Title: Cooperation between Intrinsically Disordered Regions regulates CBP condensate behaviour.

Authors:

Katie L. Gelder[1,4,5,6], Nicola A. Carruthers[1,4,5,6], Sophie Ball[1,8], Mark Dunning[7], Timothy D. Craggs[3,4,5], Alison E. Twelvetrees[2,4,5,6], Daniel A. Bose[1,4,5,6].

Affiliations

1. Molecular and Cellular Biology, School of Biosciences, Faculty of Science, The University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, UK.
2. Division of Neuroscience, The School of Medicine and Population Health, Faculty of Health, The University of Sheffield, Glossop Road, Sheffield, S10 2HQ, UK.
3. Department of Chemistry, Faculty of Science, The University of Sheffield, Dainton Building, Brook Hill, S3 7HF, UK.
4. Single Molecules at Sheffield (SM@SH), The University of Sheffield.
5. Nucleic Acids Institute, The University of Sheffield.
6. Neuroscience Institute, The University of Sheffield.
7. Sheffield Bioinformatics Core, Faculty of Health, The University of Sheffield.
8. Department of Immunology and Inflammation, Imperial College London.

Corresponding author: Daniel Bose, e-mail: d.bose@sheffield.ac.uk

Keywords

CBP/p300, gene regulation, biomolecular condensates, transcription, coactivators, enhancers, live-cell imaging, endogenous tagging.

Character count: 105368

Abstract

Creb binding protein (CBP) and *EP300* (p300) are transcription coactivators that bind at *cis*-regulatory elements and regulate gene expression in multicellular organisms. They contain large intrinsically disordered regions (IDRs), and form biomolecular condensates that are important for transcription regulation. However, little is known about how different IDRs work together to contribute to the behaviour of CBP condensates, knowledge which is essential to fully understand their biological function. Here, we show that endogenous CBP forms fluid-like condensates in the nucleus, and that properties of CBP condensates depend on IDRs dispersed throughout its sequence. Adjacent IDRs in the C-terminus form a regulatory module that balances positive and negative contributions, to ensure that CBP condensates are responsive to other IDRs and to post translational lysine acetylation. These distinct behaviours also change the characteristics of endogenous CBP condensates, including changing patterns of chromatin localisation and altering behaviours in live cell imaging. Cooperation between different IDRs therefore ensures that CBP condensates are poised to respond to their environment, allowing CBP to effectively integrate signalling events and control diverse transcriptional outputs.

Introduction

Enhancers are non-coding regions of DNA in the genome, which orchestrate developmental or environmental gene expression by recruiting transcription factors (TFs). Their activity relies on coactivators, such as the histone acetyltransferases (HATs) *Creb binding protein* (CBP) and its paralog *EP300* (p300), which integrate TF binding and control levels of target gene transcription by finely balancing lysine acetylation and assembly of regulatory protein complexes. CBP and p300 dependent lysine acetylation, including on histones (e.g H3K27ac, H3K18ac and H3K122ac), and multiple transcription regulatory proteins, are characteristic of regulatory elements and important for regulating transcription and maintaining transcription hubs [1–9]. Catalytic inhibitors of CBP and p300 cause rapid deactivation of enhancers, impair formation of the pre-initiation complex at promoters and increase chromatin association of p300 [10,11]. Despite this, PTMs that are solely due to CBP/p300 activity, such as H3K27ac[5], do not always correlate strongly with enhancer activity [12,13] and also depend on underlying transcription by RNA polymerase II (PolII) [14]. Moreover, CBP and p300 can play a structural role, contributing to PolII recruitment and maintenance of promoter-proximal pausing [8,15,16]. Large IDRs throughout the length of CBP and p300 are integral to these processes. CBP interacts with over 400 different proteins via a combination of TF interacting domains and IDRs, and is almost universally recruited to enhancers across the genome via these interactions [1,9,17,18].

The dissonance between CBP recruitment, CBP-dependent histone acetylation and enhancer activity reflects the complexity inherent in regulating CBP function. CBP activity can be fine-tuned by binding to enhancer RNAs (eRNAs) [19,20], and p300's catalytic activity is sensitive to its local environment, being downregulated by viscogens and when the catalytic HAT domain forms phase-separated condensates [21,22]. Conversely, TF binding to different regions in CBP can stimulate or inhibit its activity[23], and condensation of CBP and p300 in response to interactions with different TFs leads to increased levels of acetylation across the genome and contributes to transcriptional bursting[9]. In combination, this suggests that the activity of CBP and p300 is exquisitely responsive to its local environment at different enhancers genome-wide, but the mechanisms behind this responsiveness are unclear.

IDRs are amino acid sequences that lack stable secondary and tertiary structures(van der Lee et al. 2014; Holehouse and Kragelund 2023). One potential role of IDRs is the ability to sense and respond to changes in their surrounding environment[24,25]. While the sequence and chemical properties of IDRs that give rise to condensates are increasingly well established[26], how combinations of IDRs with distinct sequence characteristics cooperate to regulate condensate properties is less well defined. Understanding these behaviours is

especially important when, as for CBP, IDR-mediated interactions with other proteins such as TFs[9,18,27] are essential for biological function.

To address this crucial question, we asked how different IDRs cooperate to regulate full length CBP condensates, by systematically characterising how IDRs affect the formation, fluid-like properties and chromatin interactions of CBP condensates in cells. We found that not all CBP-IDRs contribute equally; instead, opposing positive and negative contributions from IDRs throughout its sequence create balance in CBP condensates, a poised state which could allow their behaviour to vary in response to different stimuli. This has important implications for the physiological behaviour of CBP in cells. One of the key roles of CBP is to integrate diverse signalling events quickly and responsively to ensure a context-specific transcriptional output; the delicately-poised nature of CBP condensates at enhancers could therefore allow combinations of TFs, RNAs and PTMs to affect CBP function by rapidly altering the properties of CBP condensates, thus contributing to unique profiles of CBP and enhancer activity.

Results

CBP forms fluid-like condensates in the nucleus.

To validate the formation of CBP condensates in the nucleus, we first imaged endogenous CBP by immunofluorescence (IF) in HEK293T cells, revealing distinct nuclear puncta containing CBP, consistent with previous observations[28–31] (Figure 1A). To uncover the behaviour of these puncta, we introduced an in-frame HaloTag to the C-terminus of endogenous CBP (CBP-Halo, Figure S1A-B). PCR from genomic DNA confirmed heterozygous insertion (Figure S1A-B) and we confirmed near-endogenous expression levels by RT-qPCR (Figure S1C) and western blotting (Figure 1B). To confirm that CBP-Halo can also form endogenous puncta, we carried out live-cell imaging using the fluorescent ligand TMR to label the HaloTag (CBP-Halo$_{TMR}$). CBP-Halo$_{TMR}$ puncta were only observable in CBP-Halo HEK293T cells upon addition TMR (Figure 1C). Importantly, CBP-Halo$_{TMR}$ condensates demonstrated hallmarks of fluid-like behaviour, as they were able to fuse and disperse over the 20s imaging period (Figure 1D, S1D & Supplementary Movie 1). We also carried out Fluorescence Recovery After Photobleaching (FRAP) experiments using Janelia Fluor® 549 Halo ligand (CBP-Halo$_{JF549}$), which showed that fluorescent signal in stained endogenous condensates recovered over a 50s timeframe following photobleaching (Figure

Figure 1

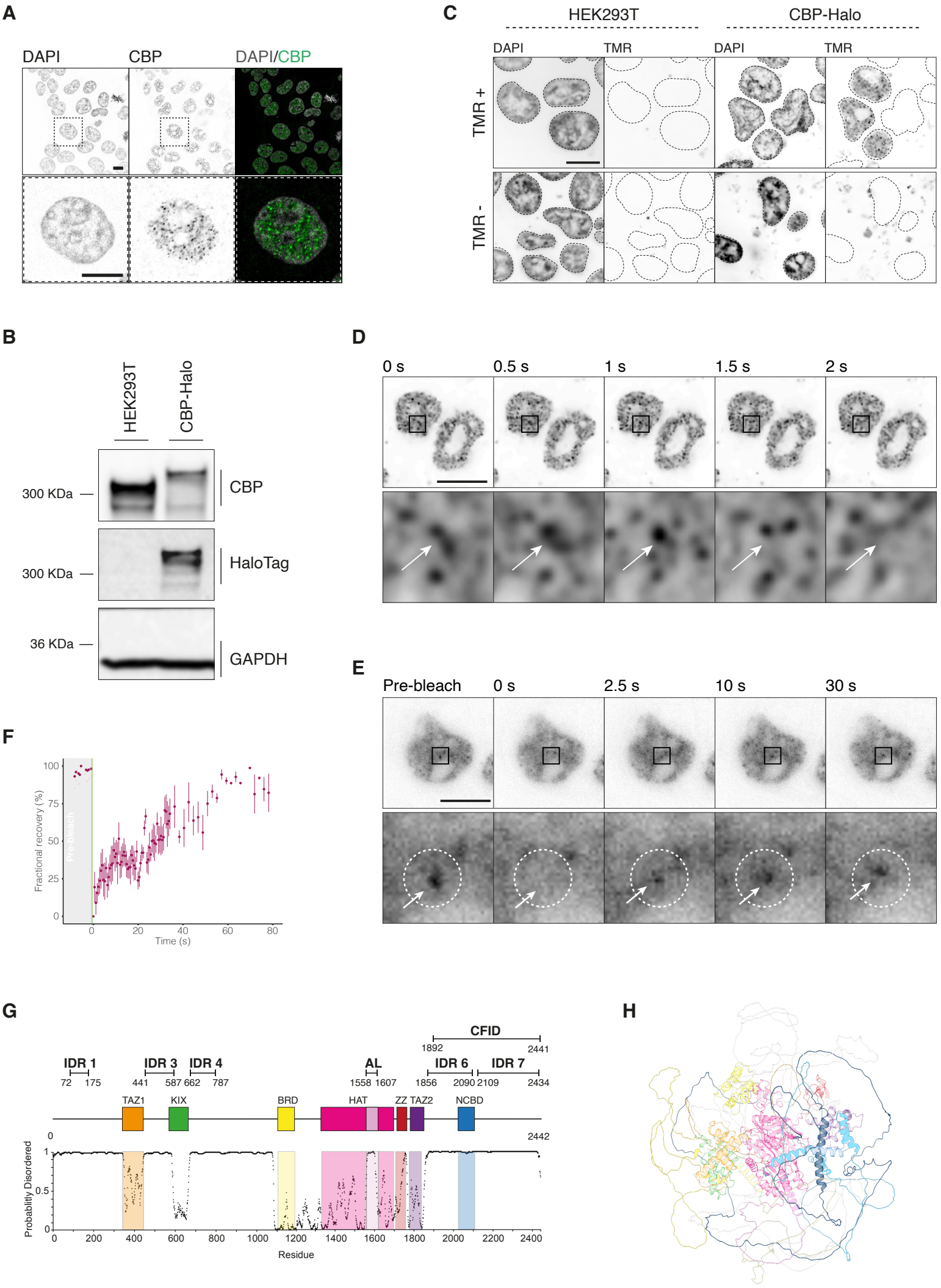

A
DAPI
CBP
DAPI/CBP
C
HEK293T
CBP-Halo
DAPI
TMR
DAPI
TMR
TMR +
TMR -
B
HEK293T
CBP-Halo
300 KDa
CBP
300 KDa
HaloTag
36 KDa
GAPDH
D
0 s
0.5 s
1 s
1.5 s
2 s
E
Pre-bleach
0 s
2.5 s
10 s
30 s
F
Fractional recovery (%)
Pre-bleach
Time (s)
G
CFID
1892
2441
IDR 1
72
175
IDR 3
441
587
IDR 4
662
787
AL
1558
1607
IDR 6
1856
2090
IDR 7
2109
2434
TAZ1
KIX
BRD
HAT
ZZ
TAZ2
NCBD
0
2442
Probability Disordered
Residue
H

Figure 1: Fluid-like behaviour of endogenous CBP condensates

A) Immunofluorescence of CBP in HEK293T cells; Scale bars: 10 µm.

B) Western blot for CBP and HaloTag in HEK293T cells and CBP-Halo HEK293T cells.

C) Confocal microscopy in fixed HEK293T and CBP-Halo HEK293T cells using the fluorescent Halo ligand TMR (CBP-Halo$_{TMR}$); Scale bars: 10 µm.

D) Still images from Supplementary Movie 1 showing the dynamic behaviour of CBP-Halo condensates. Boxes indicate regions enlarged in the lower panel; arrows highlight representative dynamic condensates; Scale bars: 10 µm.

E-F) FRAP analysis of CBP-Halo condensates. E) Still images of live cell confocal microscopy showing FRAP of endogenous CBP-Halo$_{JF549}$. Scale bars: 10 µm; F) FRAP curves for endogenous CBP-Halo$_{JF549}$. Data was collected for 32 individual nuclei, and FRAP curves were calculated for 7 condensates where the bleached area did not diffuse away from the focal plane during the recovery period. Grey shading highlights imaging pre-bleach; Green line denotes bleaching. Still images from Supplementary Movie 2

G) Domain map and predicted intrinsic disorder (DISOPRED3)[82] of full length CBP$_{wt}$. Positions of CBP-IDRs (CBP$_{IDR1-7}$, CBP$_{AL}$ and CBP$_{CFID}$) with their domain boundaries, and structured CBP domains are highlighted.

H) AlphaFold v2.3.2[32] prediction for CBP$_{wt}$. Domains are coloured the same as Figure 1G.

1E-F & Supplementary Movie 2). Together, these data clearly establish that endogenous CBP condensates in their native nuclear context have fluid-like behaviour.

IDRs in CBP have distinct abilities to form condensates.

Full length CBP_{wt} is predicted to be ~65% disordered, with these IDRs dispersed throughout its sequence (Figure 1G)[18]. This is clearly illustrated by the AlphaFold prediction of CBP_{wt} (Figure 1H)[32], which reveals structured domains interspersed with large disordered regions. In total, we identified seven IDRs within CBP for further analysis (Figure 1G & S1E). We examined the activity of five out of seven IDRs ($CBP_{IDR1-IDR7}$) identified by PONDR[33] using a threshold of 50 consecutive amino acids with a predicted disorder >50%[34]. We also focused on a C-terminal region (residues 1892-2441) that contains two of the identified IDRs (CBP_{IDR6} and CBP_{IDR7}) and the Nuclear coactivator binding domain (NCBD), which folds when interacting with TF binding partners[35,36]; as this region also interacts with Fused-in-sarcoma (FUS), we hereafter refer to it as the CBP:FUS interaction domain (CBP_{CFID}) [9,37–40]. Finally, we included the Autoregulatory Loop (AL, residues 1558-1607), a disordered regulatory region important for condensate formation of the p300 HAT domain and for RNA binding (Figure 1G)[19,22,41,42]. The selected IDRs are characterised by distinct amino acid compositions (Figure S1G) and sequence properties, including different fraction of charged residues (FCR) and net charge per residue (NCPR) [43] (Figure S1G). As the formation of biomolecular condensates in proteins relies on low complexity IDRs, and condensate properties vary according to their amino acid composition [44], we tested the different contribution of CBP-IDRs to the formation of CBP condensates.

To determine how IDRs affect CBP condensate formation, we tested the ability of each CBP-IDR to form condensates in the nucleus of HEK293T cells using the OptoDroplet system (Figure 2A)[34,38]. Candidate IDRs and controls were fused to the photoactivatable Photolyase homology region of Cry2 (Cry2PHR) to promote self-association upon exposure to blue light, together with an SV40 nuclear localization sequence (NLS) to control phase transitions in the nucleus (Figure 2B). Upon blue light exposure, CBP_{IDR7}, CBP_{CFID} and positive control FUS-IDR formed clear nuclear puncta, but surprisingly the majority of CBP-IDRs failed to form puncta, displaying a diffuse localisation similar to negative control Opto-NLS (Figure 2C & S2A). As the formation of condensates increases the variance in the grey values for a defined nucleus, this is reflected in an increase in the standard deviation. For example, a diffuse signal should have a low standard deviation, while strong changes in intensity due to condensate formation should have a high standard deviation. To quantify the

Figure 2

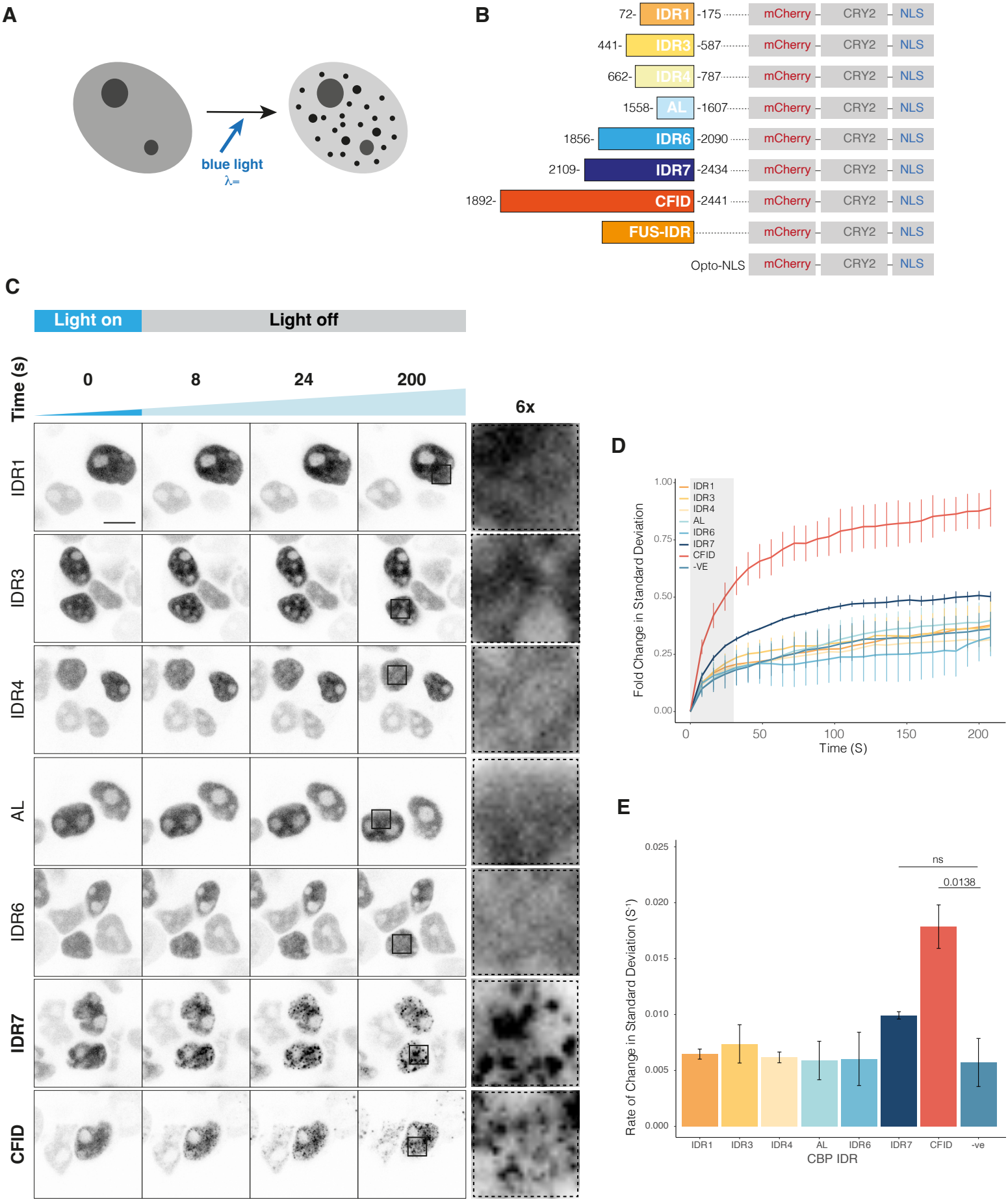

A
blue light
λ=
B
72- IDR1 -175 mCherry – CRY2 – NLS
441- IDR3 -587 mCherry – CRY2 – NLS
662- IDR4 -787 mCherry – CRY2 – NLS
1558- AL -1607 mCherry – CRY2 – NLS
1856- IDR6 -2090 mCherry – CRY2 – NLS
2109- IDR7 -2434 mCherry – CRY2 – NLS
1892- CFID -2441 mCherry – CRY2 – NLS
FUS-IDR mCherry – CRY2 – NLS
Opto-NLS mCherry – CRY2 – NLS
C
Light on
Light off
Time (s)
0
8
24
200
6x
IDR1
IDR3
IDR4
AL
IDR6
IDR7
CFID
D
Fold Change in Standard Deviation
Time (S)
IDR1
IDR3
IDR4
AL
IDR6
IDR7
CFID
-VE
E
Rate of Change in Standard Deviation (S-1)
CBP IDR
ns
0.0138
IDR1
IDR3
IDR4
AL
IDR6
IDR7
CFID
-ve

Figure 2: CBP IDRs have distinct properties and behaviours

A) Behaviour expected for condensate-forming IDRs in the OptoDroplet assay.

B) CBP-IDR constructs used for OptoDroplet assay.

C) Live-cell confocal microscopy images showing behaviour of CBP-IDRs in OptoDroplet assay. Cells were imaged for 200s following exposure to blue light. Scale bars: 10 μm

D) Fold change in the calculated standard deviation of the grey values for defined individual nuclei following blue light stimulation. Diffuse signal should have a low standard deviation, while strong changes in intensity due to condensate formation should have a high standard deviation. Grey shading highlights the 30 second window used to calculate the initial rate of change. Bars represent mean ± s.e.m, n=3.

E) Initial rate of change over the first 30 seconds following blue light exposure. AL, Autoregulatory Loop; -ve, Opto-NLS negative control. Bars represent mean ± s.e.m; significant p values are reported (ns, not significant), n=3.

observed differences in condensate forming behaviour, we therefore calculated the standard deviation of the grey values for individual nuclei following blue light stimulation. Notably, most tested CBP-IDRs behaved similarly to the negative control, with the exception of CBP_{IDR7} and CBP_{CFID} which both demonstrated a rapid increase in standard deviation following exposure to blue light (Figure 2D & S2B). We then calculated the initial rate of droplet formation over the first 30 seconds. Again, only CBP_{IDR7} and CBP_{CFID} demonstrated a greater propensity to form condensates than the negative control Opto-NLS (Figure 2C-E & S2A-D). The results suggest that different CBP-IDRs have pronounced differences in their ability to independently undergo phase transitions in the nucleus, but that in many cases their propensity to form condensates in isolation is weak. However, we also identified strong condensate forming properties in CBP_{IDR7} and CBP_{CFID} which are both located in the disordered C-terminal domain of CBP.

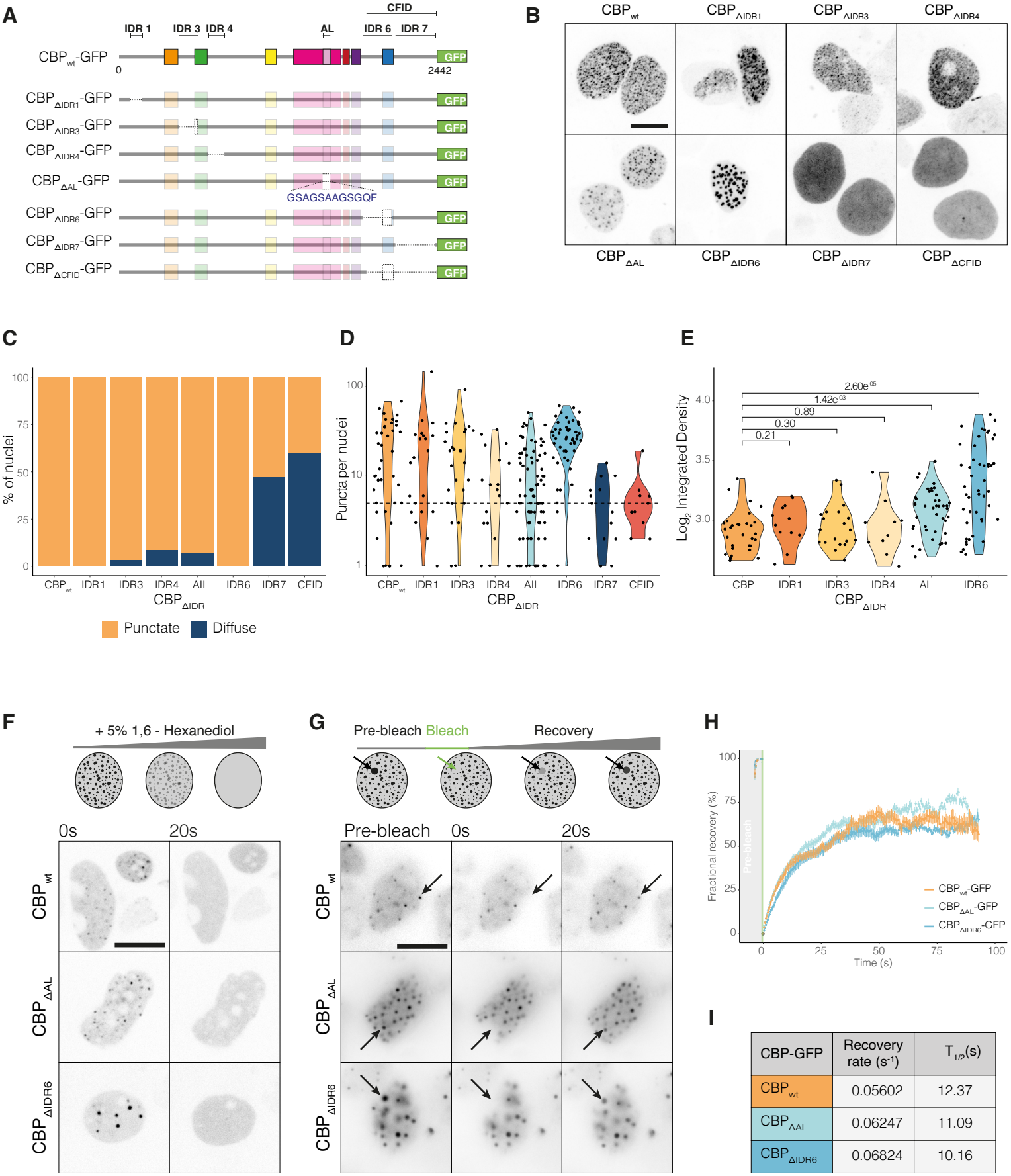
CBP-IDRs regulate the number of CBP condensates

Having established that CBP-IDRs have distinct condensate forming capabilities, we next tested how they contribute to formation of full-length CBP (CBP_{wt}) condensates. We deleted each of our candidate IDRs in CBP_{wt} fused to an in-frame C-terminal GFP-tag (CBP_{wt}-GFP, Figure 3A). To generate $CBP_{\Delta AL}$-GFP, we replaced the IDR with a flexible 12 residue linker to maintain correct folding of the HAT domain (Figure 3A) [45]. HEK293T cells transfected with CBP-GFP constructs were fixed and imaged by confocal microscopy for quantification of nuclear puncta (Figure 3B & S3A).

Data were initially selected for successful transfection based on the presence or absence of GFP signal, then the data was blinded and scored on whether the nuclei displayed a punctate or diffuse signal (Figure 3C). When transfected with CBP_{wt}-GFP, $CBP_{\Delta IDR1}$-GFP or $CBP_{\Delta IDR6}$-GFP, 100% of scored nuclei demonstrated a punctate GFP signal (Figure 3C). Indeed, transfection with the majority of CBP-GFP constructs resulted in punctate GFP signal, as $CBP_{\Delta IDR3}$-GFP, $CBP_{\Delta IDR4}$-GFP and $CBP_{\Delta AL}$-GFP were all >91% punctate (Figure 3C). $CBP_{\Delta IDR7}$-GFP and CBP_{CFID}-GFP were exceptions, with only 52.9% and 40.0% of nuclei showing a punctate signal. However, this coarse-grained approach was insufficient to describe the full spectrum of condensate-forming behaviours exhibited by $CBP_{\Delta IDR}$-GFP.

To examine these behavioural differences in more detail, we first filtered transfected nuclei to remove oversaturated nuclei, noise and overlapping puncta (Figure S3B-D). Next, we calculated the mean number of condensates per nuclei for each CBP-GFP construct (Figure

Figure 3



I

CBP-GFP	Recovery rate (s^{-1})	$T_{1/2}$(s)
CBP$_{wt}$	0.05602	12.37
CBP$_{\Delta AL}$	0.06247	11.09
CBP$_{\Delta IDR6}$	0.06824	10.16

Figure 3: CBP-IDRs affect the behaviour of full length CBP_{wt}

A) Schematic of $CBP_{\Delta IDR}$-GFP constructs. Deleted regions are highlighted with dashed lines. For $CBP_{\Delta AL}$-GFP, the AL is replaced with a flexible 12 residue linker (blue text).

B) Confocal microscopy in fixed transfected HEK293T cells showing the behaviour of CBP-GFP condensates. Scale bars: 10 µm.

C) Percentage of nuclei displaying diffuse or punctate signal. Images were initially selected for successful transfection based on the presence or absence of GFP signal then the data was blinded and scored on whether GFP signal was punctate or diffuse.

D) Mean number of condensates per nuclei. Dashed line represents a threshold value of 6 puncta per nuclei. Data points highlight individual nuclei from 3 biological replicates.

E) Integrated intensity density of condensates formed by $CBP_{\Delta IDR}$-GFP constructs. p values were calculated using a Kruskal-Wallis test.

F) Effect of 1,6 Hexanediol on CBP-GFP puncta. HEK293T cells were transfected with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta AL}$-GFP. Images were taken at fixed time points following application of 5% w/v (final concentration) 1,6 Hexanediol. Scale bars: 10 µm. Still images from Supplementary Movies 3-5.

G-I) FRAP analysis of individual CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta AL}$-GFP puncta. G) Confocal microscopy images showing photobleaching and recovery of CBP-GFP signal. Arrow indicates the bleached puncta. Scale bars: 10 µm; H) Relative fluorescence recovery normalised to pre-bleached intensity for a 90s time period following bleaching. I) Calculated values for recovery rate over the time course following bleaching and the time to reach 50% pre-bleach intensity ($T_{1/2}$). Grey shading highlights imaging pre-bleach; Green line denotes bleaching. Still images from Supplementary Movies 6-8.

3D). CBP_{wt}-GFP formed well defined condensates (average 21.1 ± 2.8 s.e.m condensates/nucleus, Figure 3B,D). The majority of CBP-IDR deletions (except $CBP_{\Delta CFID}$-GFP and $CBP_{\Delta IDR7}$-GFP) also demonstrated the ability to form condensates, as >80% of nuclei displayed a punctate signal (Figure 3C) and nuclei that passed the filter formed >6 condensates per nucleus (Figure 3D). $CBP_{\Delta IDR1}$-GFP (22.8 ± 6.8 s.e.m condensates/nucleus) and $CBP_{\Delta IDR3}$-GFP (19.9 ± 3.7 s.e.m condensates/nucleus) behaved similarly to CBP_{wt}-GFP, while $CBP_{\Delta IDR4}$-GFP (9.8 ± 2.2 s.e.m condensates/nucleus) and $CBP_{\Delta AL}$-GFP (10.9 ± 1.2 s.e.m condensates/nucleus) showed a ~50% decrease in the average number of condensates compared to CBP_{wt}-GFP (Figure 3D). Interestingly, $CBP_{\Delta IDR6}$-GFP formed ~30% more condensates (27.6 ± 1.9 s.e.m condensates/nucleus) than CBP_{wt}-GFP (Figure 3B,D).

In contrast, 60% of $CBP_{\Delta CFID}$-GFP and 47.1% of $CBP_{\Delta IDR7}$-GFP nuclei displayed a diffuse GFP signal (Figure 3C) and a ~75% decrease in quantifiable condensates per nucleus compared to CBP_{wt}-GFP ($CBP_{\Delta CFID}$-GFP, 5.6 ± 1.4 s.e.m condensates/nucleus; $CBP_{\Delta IDR7}$-GFP, 5.3 ± 1.1 s.e.m condensates/nucleus (Figure 3B,D). Visual inspection of nuclei showed that unlike other CBP-GFP constructs, any residual condensates which passed the unbiased filter were not clearly defined puncta (Figure 3B). We were therefore confident that both $CBP_{\Delta CFID}$-GFP and $CBP_{\Delta IDR7}$-GFP failed to form condensates. Together with the ability of CBP_{CFID} and CBP_{IDR7} to form condensates in the OptoDroplet experiments (Figure 2C-E), the result emphasises the importance of these C-terminal IDRs for promoting CBP_{wt} condensate formation.

Unexpectedly, the CBP-IDR deletions also affected the variability of condensate formation. For CBP_{wt}-GFP, the standard deviation (SD) in the number of condensates formed per nucleus was 86% of the mean (SD = 18.3, mean = 21.1). The majority of CBP-IDR deletions displayed a similarly high ratio of SD to mean number of condensates, ranging from 137% for $CBP_{\Delta IDR1}$-GFP to 75% for $CBP_{\Delta IDR7}$-GFP. The exception was again $CBP_{\Delta IDR6}$-GFP, where the SD was only 49% of the mean (SD = 13.6, mean = 27.6; Figure 3D). The result suggests that as well as affecting the overall number of condensates, different IDRs can also affect the variability in condensate formation, and suggests that CBP_{IDR6} could help to facilitate different condensate forming behaviours by promoting variability in the number of CBP condensates.

CBP-IDRs regulate the size and intensity of CBP condensates

As well as the number of condensates per nucleus, deletion of CBP-IDRs also altered the size and intensity of CBP condensates (Figure 3B). To better understand this behaviour, we calculated the integrated intensity density of puncta for condensate-forming $CBP_{\Delta IDR}$-GFP constructs (Figure 3E). CBP_{wt}-GFP condensates had relatively uniform size and intensity (mean IID = 870 ± 72 s.e.m.), and most $CBP_{\Delta IDR}$-GFP constructs again behaved similarly to CBP_{wt}-GFP ($CBP_{\Delta IDR1}$-GFP, mean IID = 991 ± 93 s.e.m.; $CBP_{\Delta IDR3}$-GFP, mean IID = 960 ± 100 s.e.m.; $CBP_{\Delta IDR4}$-GFP, mean IID = 966 ± 178 s.e.m.). The two exceptions were $CBP_{\Delta AL}$-GFP which demonstrated a 1.5-fold increase (mean IID = 1379 ± 105 s.e.m.) and $CBP_{\Delta IDR6}$-GFP which demonstrated a 3-fold increase (mean IID = 2647 ± 281 s.e.m.) in integrated density compared with CBP_{wt}-GFP (mean IID = 870 ± 72 s.e.m.) (Figure 3E). The results suggest that CBP_{IDR6} and CBP_{AL} may restrict the size and intensity of full length CBP condensates.

To test whether hydrophobic interactions contribute to the maintenance of CBP condensates, we treated cells transfected with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta AL}$-GFP with 5% 1,6-hexanediol (1,6-HD), an aliphatic alcohol that disrupts weak hydrophobic interactions[46]. Condensates formed by all three constructs dissolved in <20s following application of 5% 1,6-HD (Figure 3F), suggesting that hydrophobic interactions are important. Next, to determine whether the larger CBP-GFP condensates behave as fluids, gels or solids in the nucleus, we carried out FRAP on individual CBP_{wt}-GFP, $CBP_{\Delta AL}$-GFP and $CBP_{\Delta IDR6}$-GFP puncta (Figure 3G-H). Following bleaching, all condensates recovered rapidly (Recovery rate: CBP_{wt}-GFP = 0.11 s^{-1}, $CBP_{\Delta AL}$-GFP = 0.12 s^{-1}, $CBP_{\Delta IDR6}$-GFP = 0.13 s^{-1}, Figure 3I). Therefore, the tested condensates all demonstrated fluid-like behaviour, rather than solid or gel-like properties, and only displayed subtle differences in nuclear mobility (Figure 3I).

Taken togther, these results highlight clear differences in the ability of different CBP-IDRs to contribute to CBP condensates and their properties. Consistent with OptoDroplet data that demonstrated their ability to form condensates, (Figure 2C-E), the C-terminal IDRs CBP_{IDR7} and CBP_{CFID} contribute strongly to promoting CBP condensate formation. In contrast, CBP_{AL} and CBP_{IDR6} play a more complex regulatory role. Their removal increases the size and intensity of condensates, but while deletion of CBP_{AL} decreases the number of condensates, deletion of CBP_{IDR6} increases their number and also reduces the variability in the number of CBP condensates formed.

CBP condensates are regulated by acetylation in the AL.

Next, we examined the effect of the AL on CBP condensates. The AL contains 13 lysine residues (Figure S1F) and is strongly enriched for charged amino acids (FCR=0.44, Figure S1G); it is essential for controlled CBP HAT activity as the AL blocks substrate binding to the active site to inhibit the acetyltransferase activity of CBP. Factors which displace the AL, such as binding to RNA [19], autoacetylation[41] or interacting with other HAT domains in *trans* [47] can stimulate the catalytic activity of CBP and p300. *In vitro*, the AL regulates condensate formation of recombinant p300 HAT domain [22], and is subject to acetylation on multiple lysines, which negatively regulates condensate formation [22,41].

In $CBP_{\Delta AL}$-GFP data, we observed a ~50% reduction in the number of condensates compared to CBP_{wt}-GFP (Figure 3D), consistent with the AL positively regulating condensate formation. However, when comparing the integrated density of condensates, $CBP_{\Delta AL}$-GFP formed significantly larger and more intense condensates than CBP_{wt}-GFP (Figure 3E), suggesting an inhibitory role. To explain these seemingly contradictory observations, we hypothesised that the inhibitory effect could result from lysine acetylation in the AL, which may function to fine tune the properties of CBP condensates, even when these are driven by other IDRs. To test this hypothesis, we generated double mutants on a background of CBP-IDR deletion mutants: $CBP_{\Delta AL}$-GFP and CBP_{KTG}-GFP, where all 13 lysine residues in the AL were mutated to glycine to prevent acetylation (Figure 4A).

To assess the properties of these double mutants, HEK293T cells transfected with CBP-GFP constructs were fixed and imaged by confocal microscopy for quantification of nuclear puncta (Figure 4B & S4A). As before, data were initially selected for successful transfection before the data was blinded and scored for whether the GFP signal was punctate or diffuse (Figure 4C). Due to their inability to form condensates, $CBP_{\Delta IDR7}$-GFP and $CBP_{\Delta CFID}$-GFP were not affected by $CBP_{\Delta AL}$-GFP and CBP_{KTG}-GFP, as >60% of nuclei remained diffuse and quantifiable cells formed <6 condensates/nucleus (Figure 4C & S4A-C).

On a CBP_{wt}-GFP background, $CBP_{\Delta AL}$-GFP and CBP_{KTG}-GFP had similar effects, causing an increase in the fraction of diffuse nuclei (CBP_{wt}-GFP = 0% diffuse; $CBP_{\Delta AL}$-GFP = 6.1% diffuse; CBP_{KTG}-GFP = 20.4% diffuse, Figure 4C). However, this effect was masked on a $CBP_{\Delta IDR6}$-GFP background, where the fraction of diffuse nuclei was restored to the same level as CBP_{wt}-GFP ($CBP_{\Delta AL}$-GFP = 0% diffuse; CBP_{KTG}-GFP = 0% diffuse, Figure 4C).

Figure 4

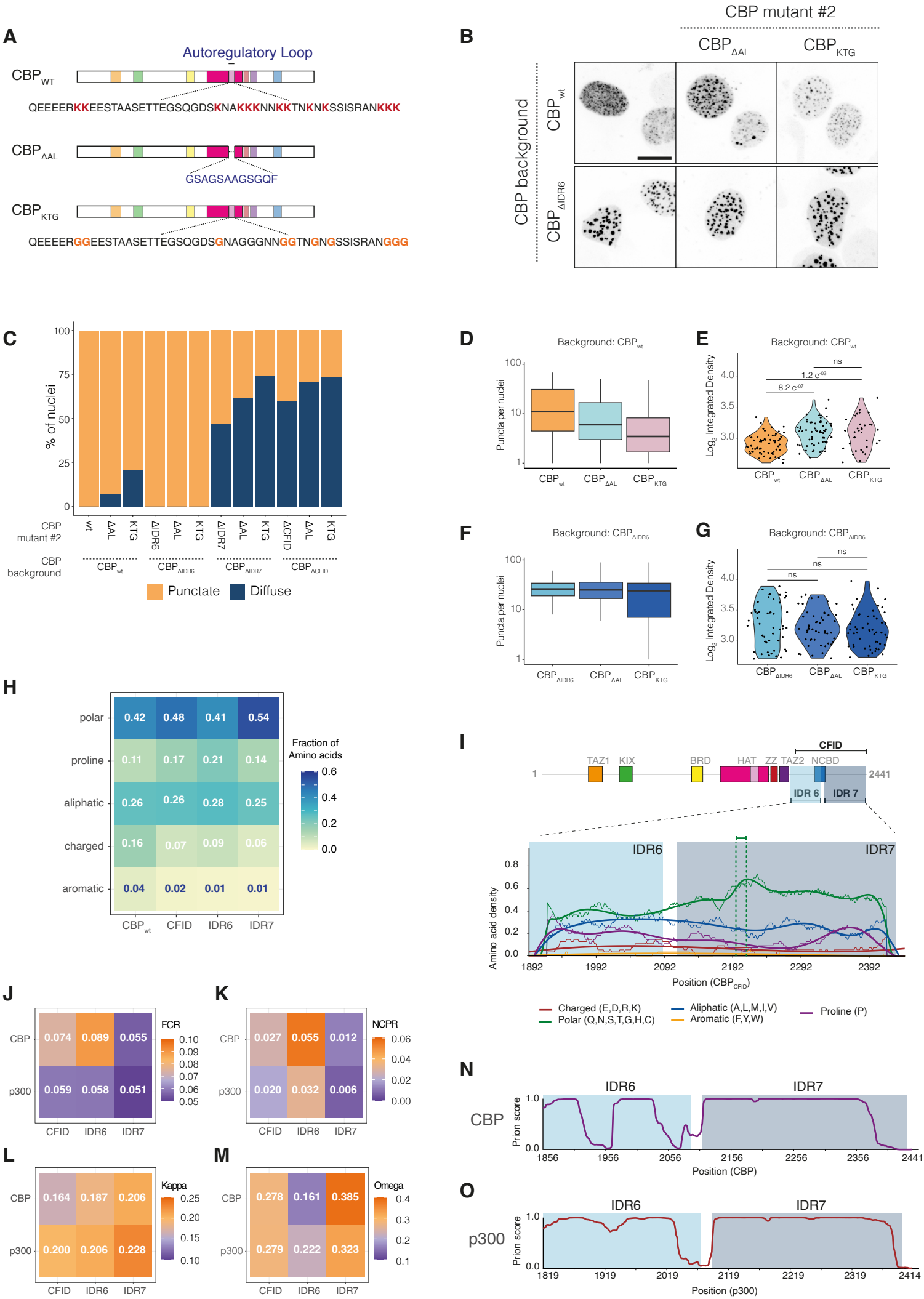

A
Autoregulatory Loop
CBP WT
QEEEERKKEESTAASETTEGSQGDSKNAKKKNNKKTNKNKSSISRANKKK
CBP ΔAL
GSAGSAAGSGQF
CBP KTG
QEEEERGGEESTAASETTEGSQGDSGNAGGGNNGGTNGNGSSISRANGGG
B
CBP mutant #2
CBP ΔAL
CBP KTG
CBP background
CBP wt
CBP ΔIDR6
C
% of nuclei
CBP mutant #2
CBP background
Punctate
Diffuse
D
Background: CBP wt
Puncta per nuclei
E
Background: CBP wt
Log2 Integrated Density
ns
1.2 e-03
8.2 e-07
F
Background: CBP ΔIDR6
G
Background: CBP ΔIDR6
H
polar
proline
aliphatic
charged
aromatic
Fraction of Amino acids
CFID
IDR6
IDR7
I
CFID
TAZ1
KIX
BRD
HAT
ZZ
TAZ2
NCBD
IDR 6
IDR 7
Amino acid density
Position (CBP CFID)
Charged (E,D,R,K)
Polar (Q,N,S,T,G,H,C)
Aliphatic (A,L,M,I,V)
Aromatic (F,Y,W)
Proline (P)
J
FCR
CBP
p300
K
NCPR
L
Kappa
M
Omega
N
Prion score
Position (CBP)
O
Position (p300)

Figure 4: Contribution of CBP_{AL} and IDR sequence properties

A) CBP_{AL} constructs showing sequences of CBP_{wt}, $CBP_{\Delta AL}$-GFP and CBP_{KTG}-GFP. In $CBP_{\Delta AL}$-GFP the deleted region was replaced by a 12-residue flexible linker (blue); Lysine residues subject to acetylation are highlighted (red, bold); positions of K>G mutations are highlighted (orange, bold).

B) Fixed cell confocal microscopy of CBP_{AL} double mutant constructs. Scale bars: 10 µm.

C) Percentage of nuclei displaying diffuse or punctate signal.

D-E) Comparison of $CBP_{\Delta AL}$-GFP and CBP_{KTG}-GFP mutant behaviour to the CBP_{wt} background: D) Mean number of condensates per nuclei; E) Integrated intensity density. p values were calculated using a Kruskal-Wallis test.

F-G) Comparison of $CBP_{\Delta AL}$-GFP and CBP_{KTG}-GFP mutant behaviour to the $CBP_{\Delta IDR6}$ background: F) Mean number of condensates per nuclei; G) Integrated intensity density. p values were calculated using a Kruskal-Wallis test.

H) Sequence properties of CBP_{wt}, CBP_{CFID}, CBP_{IDR6} and CBP_{IDR7}. Amino acids are grouped by property: Polar (Q,N,S,T,G,H,C); Proline (P); Aliphatic (A,L,M,I,V); Charged (E,D,R,K); Aromatic (F,Y,W) based on localCIDER defaults[43]. Numbers indicate the fraction of amino acids in each domain with the grouped property.

I) Linear amino acid composition of CBP_{CFID} calculated using localCIDER[43]. The relative positions of CBP_{IDR6} (light blue) and CBP_{IDR7} (Grey) are highlighted, along a schematic of their positions in the context of CBP_{wt} and CBP_{CFID}; The position of the 18 residue polyQ expansion is highlighted by the green bar and dashed lines. Amino acids are grouped by property: Charged (magenta); Polar (green); Aliphatic (blue); Aromatic (yellow); Proline (purple); Thin lines give the per-residue composition; Sliding window used to calculate local composition=50 residues; Stepsize between windows = 1 residue.

J-M) Comparison of sequence properties for CBP_{CFID}, CBP_{IDR6} and CBP_{IDR7} (Top panel) and equivalent regions in corresponding domains of p300 (bottom panel). J) Fraction of Charged Residues (FCR); K) Net Charge Per Residue (NCPR); L) Linear mixing of charged residues (κ)[49]; M) Linear mixing of charged residues and proline (Ω)[48]. All values were calculated using LocalCIDER[43].

N-O) PLAAC analysis[50,51] of CBP (purple) and p300 (Magenta) C-terminal regions. The relative positions of CBP_{IDR6} and CBP_{IDR7} are highlighted Positions in p300 are based on local

pairwise sequence alignment using EMBOSS Matcher[85](Full alignments are available in Supplementary material).

$CBP_{\Delta AL}$-GFP and CBP_{KTG}-GFP also reduced the number of CBP condensates per nucleus compared to CBP_{wt}-GFP (CBP_{wt}-GFP, 18.7 ± 1.9 s.e.m. condensates/nucleus; $CBP_{\Delta AL}$-GFP, 10.9 ± 1.2 s.e.m. condensates/nucleus; CBP_{KTG}-GFP, 7.3 ± 1.2 s.e.m. condensates/nucleus, Figure 4D), but generated a ~1.5-fold increase in their size and intensity (CBP_{wt}-GFP, mean IID = 865 ± 47 s.e.m.; $CBP_{\Delta AL}$-GFP, mean IID = 1379 ± 105 s.e.m.; CBP_{KTG}-GFP, mean IID = 1370 ± 146 s.e.m., Figure 4E). As the AL in the CBP_{KTG}-GFP construct is intact, but incapable of being acetylated, the similar effect of each mutant suggests that lysine acetylation in CBP_{wt}-GFP could limit the size of CBP condensates.

In contrast to CBP_{wt}, deletion of CBP_{IDR6} reduced the impact of $CBP_{\Delta AL}$-GFP and CBP_{KTG}-GFP on condensate formation, as double mutants only displayed subtle changes in the number of condensates formed ($CBP_{\Delta IDR6}$-GFP, 27.6 ± 1.9 s.e.m. condensates/nucleus; $CBP_{\Delta IDR6/\Delta AL}$-GFP, 27.5 ± 2.3 s.e.m. condensates/nucleus; $CBP_{\Delta IDR6/KTG}$-GFP, 23.0 ± 2.2 s.e.m. condensates/nucleus, Figure 4F). Moreover, although the size and intensity of condensates formed by double mutants was ~25% reduced compared with $CBP_{\Delta IDR6}$-GFP alone, this was still ~2.5-fold greater than for CBP_{wt}-GFP ($CBP_{\Delta IDR6}$-GFP, mean IID = 2647 ± 281 s.e.m.; $CBP_{\Delta IDR6/\Delta AL}$-GFP, mean IID = 2081 ± 181 s.e.m.; $CBP_{\Delta IDR6/KTG}$-GFP, mean IID = 1976 ± 243 s.e.m., Figure 4G), reflecting the stronger ability of $CBP_{\Delta IDR6}$-GFP to form condensates.

A similar effect was observed for the variability in the condensate number. While $CBP_{\Delta IDR6}$-GFP single mutants produced a large decrease in the ratio compared to CBP_{wt}-GFP (CBP_{wt}-GFP = 86%; $CBP_{\Delta IDR6}$-GFP = 49% Figure 3D), the decrease was smaller in the $CBP_{\Delta IDR6/\Delta AL}$-GFP and $CBP_{\Delta IDR6/KTG}$-GFP double mutants ($CBP_{\Delta IDR6/\Delta AL}$-GFP, = 64%; $CBP_{\Delta IDR6/KTG}$-GFP = 76%). Again, this reflects the stronger ability of $CBP_{\Delta IDR6}$-GFP to form condensates, and supports a role for CBP_{IDR6} as a promoter of variability in CBP condensates.

Taken together, these results suggest that the negative regulation of CBP condensates due to lysine acetylation in the CBP_{AL} is masked by the positive effect caused by deletion of CBP_{IDR6} and suggest that $CBP_{\Delta IDR6}$ can help facilitate variation in condensate behaviours. They therefore imply that an intact CBP_{IDR6} is important for allowing more subtle regulation of condensate properties by other IDRs to occur. The data highlights how the properties of full-length CBP_{wt} condensates are maintained by hierarchical cooperation between multiple IDRs, with the strongest contributions overriding weaker contributions to affect the overall condensate balance.

Opposing properties of CBP-IDRs reflect their sequence characteristics

The sequence properties of IDRs confer different behaviours on the ensemble of configurations adopted by that IDR. To investigate whether their amino acid sequence could explain the opposing abilities of CBP_{IDR6} and CBP_{IDR7} to regulate condensate formation, we used localCIDER[43] to examine the sequence properties of CBP_{CFID}, alongside CBP_{IDR6} and CBP_{IDR7} in isolation (Figure 4G). Compared to CBP_{wt}, CBP_{CFID} is enriched for residues that promote expanded chain conformations (Polar amino acids and proline)[48], but depleted for charged and aromatic residues. However, there is more variation in sequence within CBP_{CFID}: For example, polar amino acids are more enriched in CBP_{IDR7} (54%) than CBP_{IDR6} (41%) and CBP_{IDR6} has a very high proportion of proline residues (21%) compared to CBP_{IDR7} (14%). The linear distribution of amino acids (Figure 4H) shows that amino acid densities are relatively evenly distributed throughout CBP_{IDR6}, but less so for CBP_{IDR7}. This is especially true for polar amino acids, which are concentrated around an 18-residue polyQ expansion (residues 2199-2216) in CBP_{IDR7}, and proline which is concentrated towards the C-terminus (Figure 4H). The divergence in amino acid composition between CBP_{IDR6} and CBP_{IDR7} is reflected in higher values of FCR for CBP_{IDR6} (FCR=0.089) than for CBP_{IDR7} (FCR=0.055) and CBP_{CFID} (FCR=0.074) (Figure 4I), and higher values of NCPR for CBP_{IDR6} (NCPR=0.055) then for CBP_{IDR7} (NCPR=0.012 and CBP_{CFID} (NCPR=0.027)(Figure 4J).

As well as their overall composition, the sequence patterning of amino acids with different properties in IDRs can also be described. For oppositely charged positive and negative residues, the parameter κ describes the clustering of charge blocks; for perfectly mixed sequences $\kappa \rightarrow 0$ and for sequences where charged residues cluster into blocks, $\kappa \rightarrow 1$[43,49]. In general, low values of κ favour the extended conformations characteristic of IDRs. C-terminal CBP-IDRs have relatively low values of κ (CBP_{IDR6}=0.187, CBP_{IDR7} =0.206 and CBP_{CFID}=0.164; Figure 4K), suggesting that oppositely charged residues in these regions of CBP are well mixed, and that they adopt extended chain conformations [49]. A second parameter, Ω, describes the linear mixing of proline and charged residues with all other residues. Similar to κ, $\Omega \rightarrow 0$ for well mixed sequences and $\Omega \rightarrow 1$ when proline and charged residues are clustered in blocks; again, low values of Ω favour extended conformations. CBP_{CFID} has a relatively low value of Ω (Ω=0.278) suggesting that these regions of CBP adopt extended chain conformations [48]. However, there is a pronounced difference in Ω between CBP_{IDR6} (Ω =0.161) and CBP_{IDR7} (Ω=0.385; (Figure 4L), potentially reflecting a more even distribution of residues in CBP_{IDR6} and the clustering of polar residues around the polyQ

region and of prolines at the C-terminus of CBP_{IDR7} (Figure 4H). While their sequence parameters suggest these regions adopt extended conformations consistent with disorder, the different properties of CBP_{IDR6} and CBP_{IDR7} relative to each other is consistent with the differing behaviour of these two IDRs in OptoDroplet experiments (Figure 2C-D) and when deleted in CBP-GFP (Figure 3B-E).

Next, as enrichment for polar residues such as glutamine and arginine, as observed for CBP_{IDR7}, is associated with prion forming domains[50], we examined the sequence of of CBP_{CFID} using a PLAAC analysis[50,51]. Consistent with the enrichment of polar residues in CBP_{IDR7}, it was predicted to form a large, contiguous prion-like domain. However, while it still was predicted to have some prion-like properties, in CBP_{IDR6} these were discontinuous (Figure 4M).

We also asked whether the sequence characteristics of these C-terminal domains were the same for CBP and its paralog p300. Over their entire sequence, CBP and p300 are 60.8% identical (70.7% similarity). While this identity is as high as 87.3% (94.4% similarity) in the HAT domain, for CBP_{CFID}, sequence identity falls to 51.7% (62.2% similarity). For CBP_{IDR6} and CBP_{IDR7}, these values are 48.1% identity (58.0% similarity) and 56.1% identity (66.2% similarity) respectively (Figure S4D & Supplementary material). This divergence in sequence identity is reflected in a divergence in their sequence properties. For the sequence parameters FCR, NCPR and Ω, the difference between CBP_{IDR6} and CBP_{IDR7} is much more pronounced than the equivalent difference between the equivalent $p300_{IDR6}$ and $p300_{IDR7}$ IDRs (Figure 4I-J, L). The opposite was true for $\varkappa$, where the difference between $p300_{IDR6}$ and $p300_{IDR7}$ was greater than the equivalent difference in CBP. Similarly, while CBP_{IDR6} and CBP_{IDR7} are predicted to have discontinuous and contiguous prion-like properties respectively (Figure 4M), PLAAC analysis[50,51] of the C-terminal IDRs of p300 demonstrated that the corresponding regions are both predicted to have largely contiguous prion-like properties (Figure 4N). Although we did not test the physiological consequences of these differences between the sequence properties of CBP and p300 here, the results suggest that the balancing of CBP condensates orchestrated by cooperation between CBP_{IDR6} and CBP_{IDR7}, could be different in p300.

Finally, we examined the role of predicted secondary structure elements on the differing behaviours of CBP_{IDR6} and CBP_{IDR7}. Although proline enrichment may reduce transient secondary structures[26], it can also define the boundaries of secondary structure elements[52]. In CBP_{IDR6}, two regions flanked by proline-rich sequences were identified as having a high propensity to form α-helices using Nuclear Magnetic Resonance (NMR)[53]. To better

understand whether these regions play a role in the observed behaviour of our $CBP_{\Delta IDR6}$-GFP mutants, and whether this could affect the relationship between CBP_{IDR6} and CBP_{IDR7}, we used AlphaFold [32] to predict the structure of CBP_{CFID} (Figure S4E). The results revealed a 27-residue α-helix in CBP_{IDR6} ($CBP_{IDR6\text{-}H2}$, residues 1951-1978) corresponding to the second helix identified by NMR[53]. Surprisingly, this α-helical region was predicted to interact with a 29-residue α-helix (residues 2187-2216) in CBP_{IDR7} ($CBP_{IDR7\text{-}H1}$) which includes the polyQ expansion, and a third α-helix formed by residues 2288-2302 in CBP_{IDR7} ($CBP_{IDR7\text{-}H2}$)(Figure S4E). The resulting helical bundle has a pLDDT confidence score of >70-90 (Figure S4F). The predicted interactions between CBP_{IDR6} and CBP_{IDR7} bridge the NCBD domain (residues 2058-2116) at its N-terminus and C-terminus (Figure S4G). NCBD exists as a molten globule that folds upon binding to TFs[35]. To determine whether the predicted inter-domain interactions could be responsible for the inhibitory role of CBP_{IDR6}, we generated mutants of CBP_{wt}-GFP with deletions of CBP_{IDR6H2} ($CBP_{\Delta IDR6\text{-}H2}$-GFP) and $CBP_{IDR7\text{-}H1}$ ($CBP_{\Delta IDR7\text{-}H1}$-GFP) and imaged in transfected HEK293T cells (Figure S4H). Interestingly, behaviours of both α-helix deletions differed from larger CBP-IDR deletions: $CBP_{\Delta IDR6\text{-}H2}$-GFP formed fewer (12.5 ± 2.2 s.e.m. condensates/nucleus) and smaller (mean IID = 1182 ± 660 s.e.m.) condensates than $CBP_{\Delta IDR6}$-GFP, and $CBP_{\Delta IDR7\text{-}H1}$-GFP formed more (12.08 ± 1.9 s.e.m. condensates/nucleus) and larger (mean IID = 761 ± 147 s.e.m.) condensates more than $CBP_{\Delta IDR7}$-GFP. However, neither α-helix deletion altered the size of condensates when compared to CBP_{wt}-GFP (Figure S4I-J), therefore it seems likely that the condensate forming properties of CBP_{IDR6} and CBP_{IDR7} arise from the properties of the whole IDR sequence rather than transient helical interactions between the α-helical domains.

Taken together, the data highlight differences in the sequence properties of CBP_{IDR6} and CBP_{IDR7}, in particular the relative abundance of polar residues and proline, and the patterning factor Ω, which could account for their differences in behaviour. Interestingly these differences are more pronounced for CBP than for similar domains in p300, suggesting the possibility that the balance of condensates maintained by cooperation between CBP_{IDR6} and CBP_{IDR7} could differ between the two proteins.

CBP IDRs mediate different patterns of chromatin localisation

CBP is recruited to chromatin across the genome. To determine whether the observed differences in the behaviour of CBP condensates applied to chromatin-associated CBP populations, we made clonal HEK293T lines stably expressing doxycycline (Dox) inducible CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and CBP_{IDR7}-GFP which demonstrated a clear increase in GFP

Figure 5

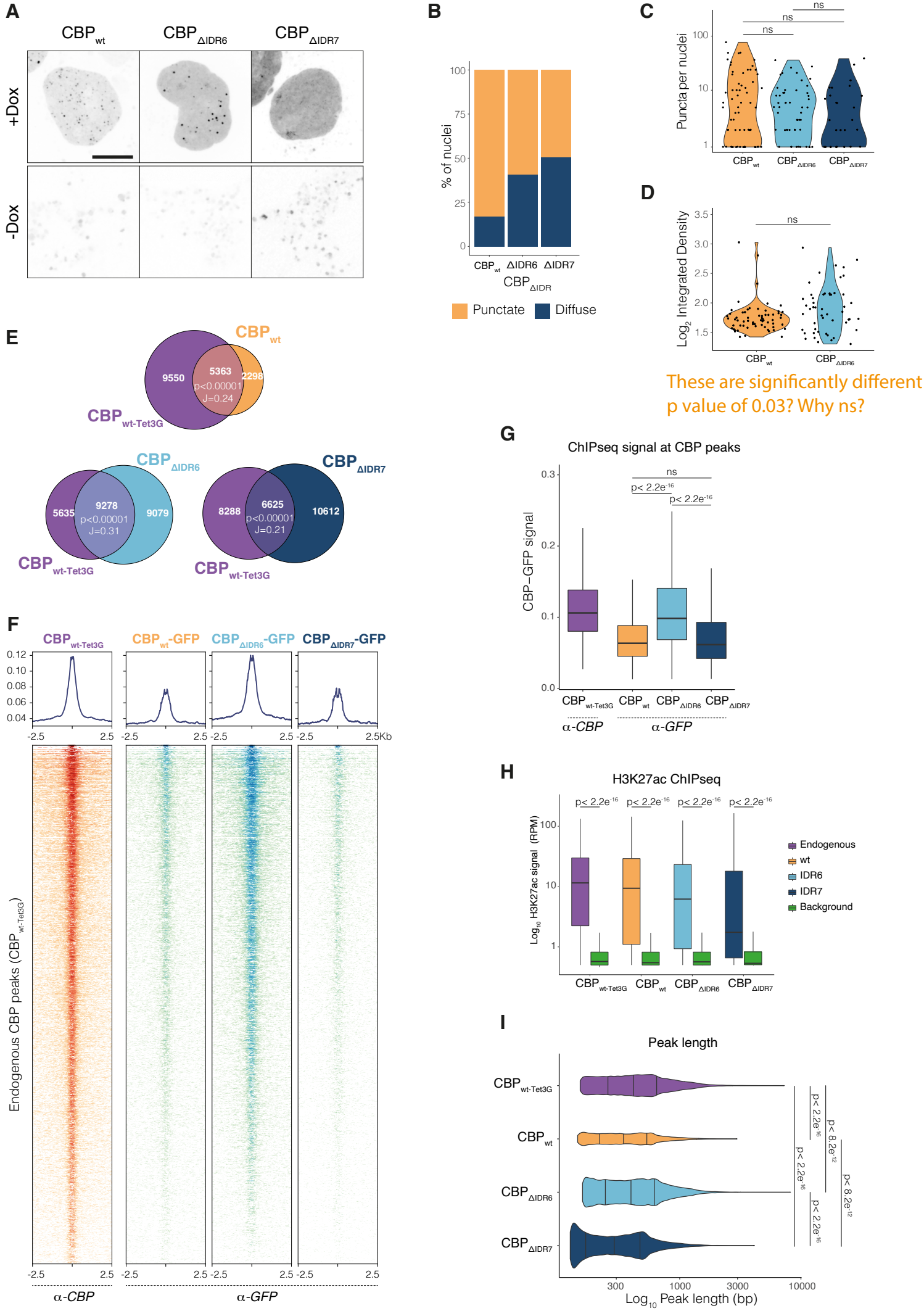

A
CBP wt
CBP ΔIDR6
CBP ΔIDR7
+Dox
-Dox
B
% of nuclei
CBP wt
ΔIDR6
ΔIDR7
CBP ΔIDR
Punctate
Diffuse
C
Puncta per nuclei
ns
D
Log2 Integrated Density
These are significantly different
p value of 0.03? Why ns?
E
CBP wt
CBP wt-Tet3G
9550
5363
2298
p<0.00001
J=0.24
CBP ΔIDR6
5635
9278
9079
p<0.00001
J=0.31
CBP ΔIDR7
8288
6625
10612
p<0.00001
J=0.21
F
CBP wt-Tet3G
CBP wt-GFP
CBP ΔIDR6-GFP
CBP ΔIDR7-GFP
Endogenous CBP peaks (CBP wt-Tet3G)
α-CBP
α-GFP
G
ChIPseq signal at CBP peaks
CBP−GFP signal
p< 2.2e-16
H
H3K27ac ChIPseq
Log10 H3K27ac signal (RPM)
Endogenous
wt
IDR6
IDR7
Background
I
Peak length
Log10 Peak length (bp)
p< 8.2e-12

Figure 5: Chromatin localisation of inducible CBP-GFP.

A) Confocal microscopy showing inducible CBP-GFP expression. Stable inducible CBP-GFP HEK293T cells were treated with Doxycycline (Dox) for 48h. Dox concentrations: CBP_{wt}-GFP, 25ngml^{-1}; $CBP_{\Delta IDR6}$-GFP, 5ngml^{-1}; and CBP_{IDR7}-GFP, 25ngml^{-1}; Scale bars: 10 µm.

B-D) Behaviour of CBP-GFP condensates following Dox induction. B) Percentage of nuclei displaying diffuse or punctate signal; C) Mean number of condensates per nuclei; D) Integrated intensity density of condensates formed by $CBP_{\Delta IDR}$-GFP constructs.

E) Venn diagrams showing overlap between endogenous CBP_{wt} in the maternal $HEK293T_{Tet3G}$ lines (purple) and: CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) and CBP_{IDR7}-GFP (dark blue). p values from permutation test with random regions restricted to TSS -40 kb; Jaccard statistic (J) measures overlap between two sets of intervals; perfect correlation = 1.

F) Heatmaps of CBP and GFP ChIP signal at endogenous CBP_{wt} peaks. ±2.5-kb window from center of CBP peak.

G) Reads at endogenous CBP_{wt} peaks. CBP ChIPseq reads in $HEK293T_{Tet3G}$ cells (purple); GFP ChIPseq reads in CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) and CBP_{IDR7}-GFP (dark blue). p values from Mann-Whitney U-test.

H) Read densities for H3K27ac (GEO: GSM1249889) [55] at CBP peaks compared with background control regions (green). CBP peaks from CBP ChIPseq in $HEK293T_{Tet3G}$ cells (purple) or GFP ChIPseq reads in CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) and CBP_{IDR7}-GFP (dark blue). Background shuffled control regions (green) were restricted to TSS -40 kb. p values from Mann-Whitney U-test.

I) Peak width distributions. Violin plots shows CBP peak widths in $HEK293T_{Tet3G}$ cells (purple), and GFP ChIPseq reads in CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) and CBP_{IDR7}-GFP (dark blue). Vertical lines indicate position of 25% quantiles. p values from Mann-Whitney U-test.

signal in response to increasing concentrations of Dox (Figure S5A). We chose Dox concentrations that induced CBP-GFP expression, while retaining the overall CBP expression level at near-endogenous levels using western blotting (Figure S5B).

Following Dox induction, CBP_{wt}-GFP and $CBP_{\Delta IDR6}$-GFP formed defined condensates, while $CBP_{\Delta IDR7}$-GFP displayed a more diffuse signal (Figure 5A). When scored, the lower level of expression of the inducible constructs caused all constructs to display a higher level of diffuse nuclei then when transiently transfected (CBP_{wt}-GFP = 83.1% punctate; $CBP_{\Delta IDR6}$-GFP = 59.2% punctate and $CBP_{\Delta IDR7}$-GFP = 49.4% punctate, Figure 3A & S5B). This was reflected in the number of condensates formed; in contrast to transient transfections, CBP_{wt}-GFP (11.2 ± 1.8 s.e.m. condensates/nucleus) formed more condensates than both $CBP_{\Delta IDR6}$-GFP (7.2 ± 1.0 s.e.m. condensates/nucleus) and $CBP_{\Delta IDR7}$-GFP (6.5 ± 1.6 s.e.m. condensates/nucleus). However, the size and intensity of condensates demonstrated patterns of behaviour that were more similar to the transiently transfected data, with condensates formed by CBP_{wt}-GFP (mean IID = 80.6 ± 18.3 s.e.m.) being smaller than $CBP_{\Delta IDR6}$-GFP (mean IID = 126.0 ± 21.5 s.e.m.).

To investigate how CBP-IDRs affect CBP chromatin localisation genome-wide, we used chromatin immunoprecipitation sequencing (ChIP-seq) for GFP following addition of Doxycycline to inducible cell lines, and for endogenous CBP_{wt} in the maternal $HEK293T_{Tet3G}$ cells ($CBP_{wt\text{-}Tet3G}$), from two biological replicates. $CBP_{wt\text{-}Tet3G}$ read densities were well correlated (0.855 < Pearson's coefficient < 0.934, Figure S5E; $0.763 < R^2 < 0.883$, Figure S5F) with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP across the genome, highlighting similar chromatin localisation profiles.

Next, we used MACS2 to call peaks in pooled replicates, revealing 14,913 binding sites for $CBP_{wt\text{-}Tet3G}$, 7,661 for CBP_{wt}-GFP, 18,357 for $CBP_{\Delta IDR6}$-GFP and 17,237 for $CBP_{\Delta IDR7}$-GFP); All called CBP-GFP peaks demonstrated statistically significant overlap with $CBP_{wt\text{-}Tet3G}$ (36% for CBP_{wt}-GFP, 62% for $CBP_{\Delta IDR6}$-GFP and 44% for $CBP_{\Delta IDR7}$-GFP; Figure 5B); CBP_{wt}-GFP peaks strongly intersected $CBP_{\Delta IDR6}$-GFP (77%) and $CBP_{\Delta IDR7}$-GFP (69%); 43% of $CBP_{\Delta IDR7}$-GFP peaks intersected $CBP_{\Delta IDR6}$-GFP (Figure S5G). Relative distance analysis[54] revealed that the intervals between peaks were closer than expected by chance, demonstrating that CBP peaks were spatially correlated with each other (Figure S5H). Finally, comparison of their genome-wide binding profiles revealed that $CBP_{wt\text{-}Tet3G}$, CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP were recruited to similar regions of the genome (Figure S5I). Taken together, the data demonstrate that CBP-GFP is recruited to chromatin at similar locations to endogenous CBP. As CBP recruitment is a marker of

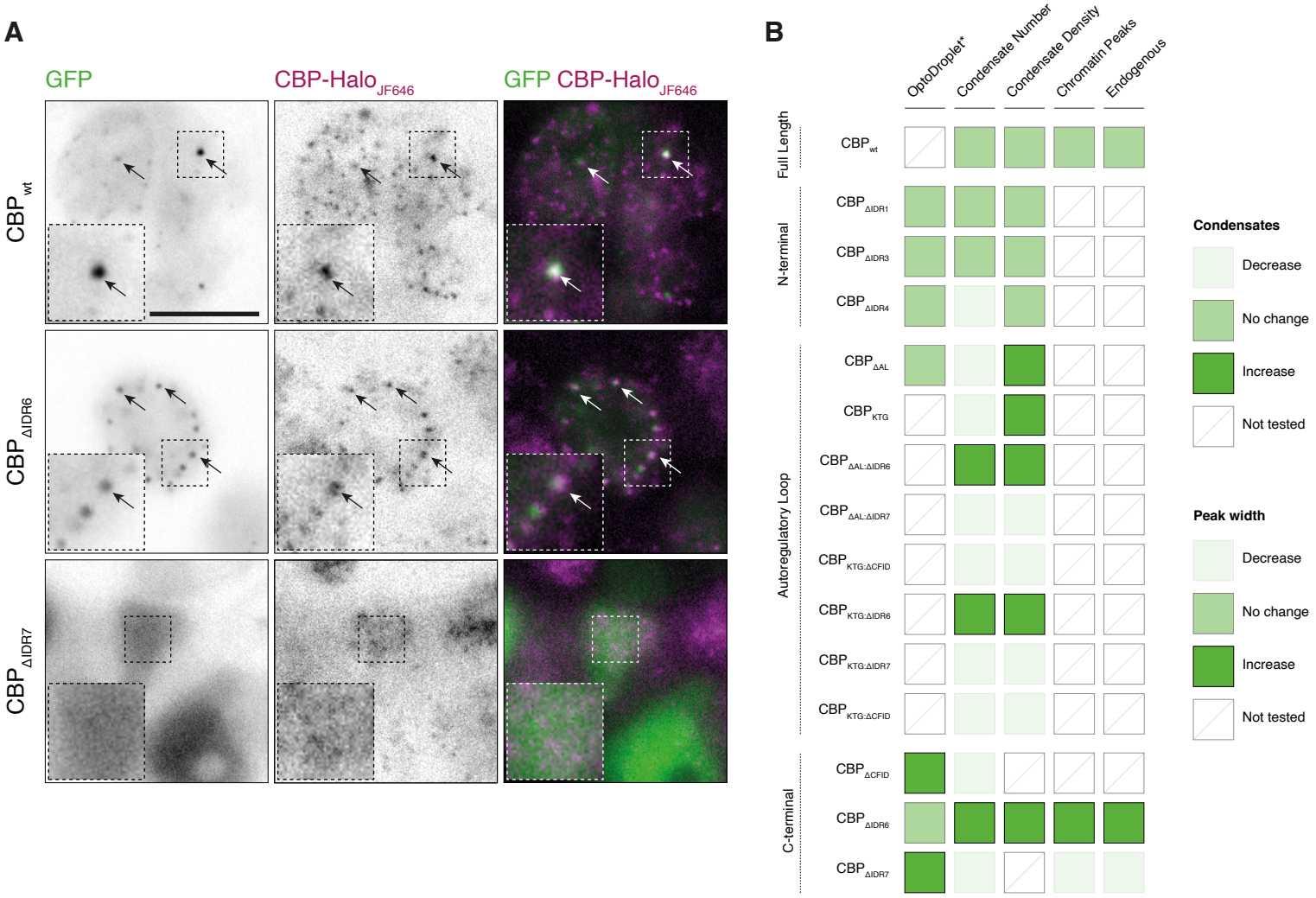
cis-regulatory elements including enhancers and promoters, we next examined whether CBP-GFP peaks bear hallmarks of regulatory chromatin. Using published ChIPseq data for H3K27ac[55], we found that CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP peaks were statistically enriched for H3K27ac (Figure 5H) compared to background controls. The data suggest that CBP-IDR mutants are recruited to chromatin at sites with features of gene regulatory elements.

Next, we asked whether CBP IDRs affect the profiles of CBP localisation. CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP reads were all enriched at endogenous CBP_{wt} peaks (Figure 5B). However, $CBP_{\Delta IDR6}$-GFP was statistically more enriched than CBP_{wt}-GFP, while there was no significant difference between the read densities of CBP_{wt}-GFP and $CBP_{\Delta IDR7}$-GFP (Figure 5D). Notably, the mean peak width (Figure 4E) of CBP_{wt}-GFP (418±3.2bp) is less than $CBP_{\Delta IDR6}$-GFP (485±2.5bp) but greater than $CBP_{\Delta IDR7}$-GFP (357±2.0 bp), although CBP_{wt} peaks were broader than all CBP-GFP peaks (519± 3bp). While we cannot eliminate the possibility that the observed differences result from variations in protein expression between the Dox inducible lines, the data strongly suggest that the behaviour of CBP-GFP constructs when chromatin-associated is consistent with the differences in behaviour observed by imaging: $CBP_{\Delta IDR6}$-GFP, which forms more and larger condensates than CBP_{wt}-GFP in live cell imaging, forms more and broader peaks on chromatin while $CBP_{\Delta IDR7}$-GFP which forms fewer condensates than CBP_{wt}-GFP in live cell imaging, forms fewer and narrower peaks on chromatin.

CBP IDRs regulate the behaviour of endogenous CBP condensates

As we showed that both endogenous CBP (Figure 1) and CBP-GFP condensates (Figure S3E-F) are fluid-like in the nucleus, we asked whether endogenous CBP_{wt} could be incorporated into the condensates formed by transfected CBP-GFP. We transfected CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP into CBP-Halo HEK293T cells, and carried out live cell confocal microscopy for CBP-GFP and CBP-Halo labelled with fluorescent JF_{646} ligand (CBP-$Halo_{JF646}$). As previously observed (Figure 3A), CBP_{wt}-GFP formed defined condensates in the nucleus (Figure 6A). When we imaged CBP-$Halo_{JF646}$ in cells transfected with CBP_{wt}-GFP, we observed more and brighter CBP-$Halo_{JF646}$ puncta than CBP_{wt}-GFP, reflecting the greater intensity of the JF_{646} fluorophore compared to GFP, which facilitated improved detection of smaller endogenous CBP-$Halo_{JF646}$ puncta (Figure 6A). Importantly, we observed no bleed-through signal in other imaging channels (Figure S6A). CBP_{wt}-GFP condensates consistently colocalized with CBP-$Halo_{JF646}$ (Figure 6A), suggesting that

Figure 6


A
GFP
CBP-Halo JF646
GFP CBP-Halo JF646
CBP wt
CBP ΔIDR6
CBP ΔIDR7
B
OptoDroplet*
Condensate Number
Condensate Density
Chromatin Peaks
Endogenous
Full Length
N-terminal
Autoregulatory Loop
C-terminal
CBP wt
CBP ΔIDR1
CBP ΔIDR3
CBP ΔIDR4
CBP ΔAL
CBP KTG
CBP ΔAL:ΔIDR6
CBP ΔAL:ΔIDR7
CBP KTG:ΔCFID
CBP KTG:ΔIDR6
CBP KTG:ΔIDR7
CBP KTG:ΔCFID
CBP ΔCFID
CBP ΔIDR6
CBP ΔIDR7
Condensates
Decrease
No change
Increase
Not tested
Peak width
Decrease
No change
Increase
Not tested

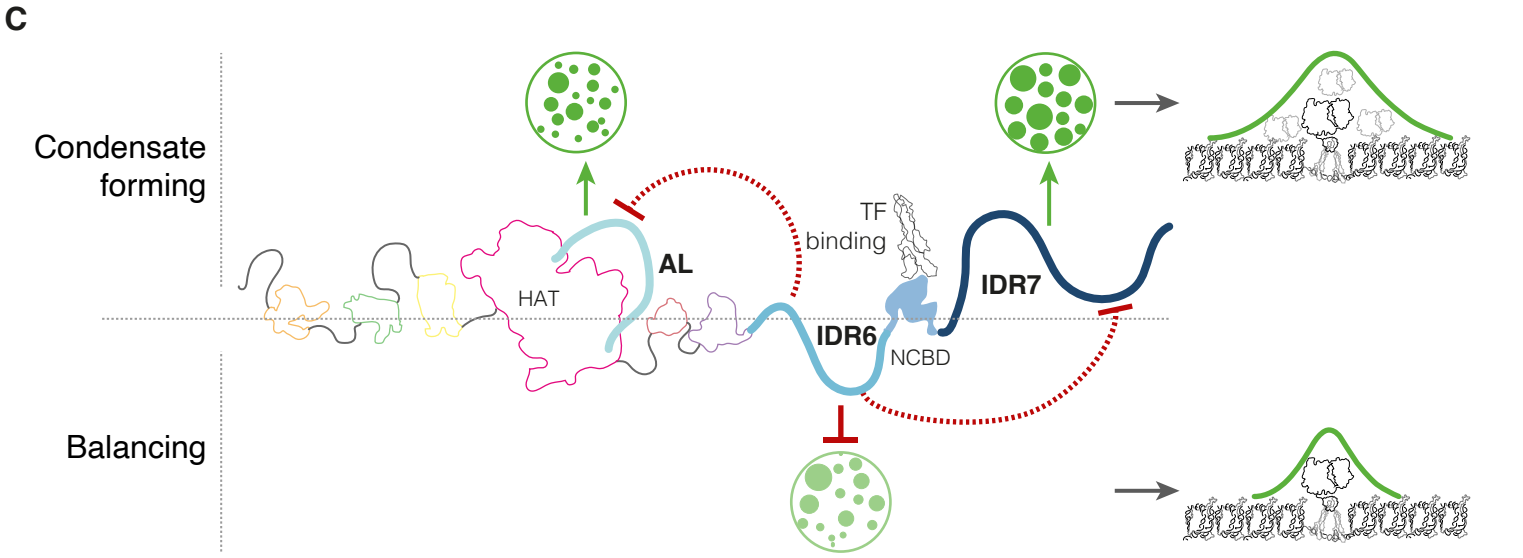

C
Condensate forming
Balancing
HAT
AL
TF binding
IDR6
NCBD
IDR7

Figure 6: $CBP_{\Delta IDR}$ mutants affect endogenous CBP condensates

A) Endogenous CBP-Halo is incorporated into CBP-GFP condensates. CBP-Halo HEK293T cells were transfected with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and CBP_{IDR7}-GFP for 24h. Endogenous CBP-Halo was labelled with Janelia Fluor® 646 Halo ligand ($CBP\text{-}Halo_{JF646}$).

B) Summary of the contribution of CBP-IDRs to CBP condensate behaviours relative to CBP_{wt}-GFP. IDRs are clustered according to whether they are: N-terminal to the HAT domain; in the Autoregulatory Loop (AL); or C-terminal to the HAT domain. All behaviours refer to behaviour of CBP following deletion of each IDR, except the OptoDroplet data which is the behaviour of the IDR in isolation. Endogenous = behaviour of endogenous CBP condensates following transfection with CBP-GFP construct.

C) Model of balancing interactions between C-terminal CBP-IDRs and their effect on CBP condensates and chromatin localisation. Positive contributions to CBP condensates made by CBP_{IDR7} and CBP_{AL} are balanced by CBP_{IDR6}; TF binding to the NCBD may alter the balance between interactions; opposing contributions to CBP condensates alter the properties of chromatin associated CBP populations.

CBP_{wt}-GFP and CBP-$Halo_{JF646}$ condensates were forming in the same locations. This pattern was replicated for both $CBP_{\Delta IDR6}$-GFP, which formed larger and more intense condensates, and $CBP_{\Delta IDR7}$-GFP, which showed a diffuse pattern of nuclear localisation. Strikingly, in both cases CBP-$Halo_{JF646}$ not only colocalized with condensates formed by transfected CBP-GFP, but also replicated their pattern of behaviour: CBP-$Halo_{JF646}$ condensates were larger and more intense when imaged with $CBP_{\Delta IDR6}$-GFP, and diffuse when imaged with $CBP_{\Delta IDR7}$-GFP. These results not only demonstrate that endogenous CBP is incorporated into overexpressed CBP-GFP condensates, but also that the patterns of condensate behaviour resulting from CBP-IDR deletion mutants are reflected in the behaviour of endogenous CBP condensates. This suggests that the deletion of CBP-IDRs can exert a dominant negative/positive effect on endogenous CBP behaviour, and implies that the behaviour of endogenous CBP condensates is also subject to regulation by opposing forces derived from different CBP-IDRs.

Our findings lead to a model where CBP condensates arise due to the combined, hierarchical contributions of IDRs dispersed throughout its sequence (summarised in Figure 6B). Not every contribution is equal: primary drivers are found in the C-terminal regions of CBP, where opposing positive and negative contributions derived from CBP_{IDR6} and CBP_{IDR7} cooperate to regulate the properties of CBP condensates, and positive contributions made by the AL are balanced by lysine acetylation (Figure 6C). CBP_{IDR6} appears to facilitate variability in condensates that allows weaker contributions from other IDRs (e.g lysine acetylation in the AL) to also impact behaviour. Notably, the expression of CBP-IDR deletion mutants can regulate the behaviour of endogenous CBP_{wt} in a dominant manner, and CBP-IDRs can modify the properties of CBP populations recruited to chromatin at regulatory regions, potentially impacting CBP function and its ability to regulate gene expression.

Discussion

CBP plays a crucial role integrating diverse signalling events and regulating distinct transcriptional responses when recruited to *cis*-regulatory elements. In order to do this, CBP must be able to integrate a diverse range of inputs and respond appropriately. Here, we have shown how IDRs dispersed throughout CBP_{wt} cooperate to maintain a balance between formation of larger and smaller condensates in the nucleus, facilitating a balance in the properties of CBP condensates that could allow them to respond to different inputs. Importantly, $CBP_{\Delta IDR}$-GFP mutants exert a dominant effect on endogenous CBP_{wt} condensates (Figure 6A), suggesting that the observed interactions between CBP-IDRs are

important for controlling the behaviour of endogenous CBP condensates. Furthermore, these behaviours are reflected in the chromatin association profiles of CBP-GFP (Figure 5F-I), suggesting that although they may be dispensable for recruitment of CBP and p300 to chromatin[11], CBP-IDRs control the formation and fluid-like properties of chromatin-bound CBP populations and thus influence diverse functions of CBP. As CBP is an important scaffolding factor at regulatory elements, this could include disrupting the recruitment and retention of other factors at regulatory elements[8,15,16].

The ability to modify the properties of CBP condensates could directly impact CBP-dependent acetylation of histone and non-histone substrates at regulatory elements. The catalytic HAT activity of CBP and p300 is downregulated by viscogens and when the HAT domain forms condensates *in vitro*[21,22]. Here we show that CBP_{IDR6} and CBP_{IDR7} form a regulatory module composed of competing positive and negative contributions to CBP condensates (Figure 6C). The intra-molecular proximity of these adjacent regions may make CBP condensates especially responsive to factor-induced changes in condensate formation, or to different post-translational modifications of CBP_{IDR6} and CBP_{IDR7}, with a corresponding impact on CBP HAT activity. Accordingly, the adenovirus protein E1A can bind to multiple sites in the N- and C-terminal tails of CBP, but is only able to inhibit the HAT activity of CBP when it interacts with a region overlapping CBP_{IDR6}; mutations in E1A which specifically disrupt its binding to CBP_{IDR6} but not other sites in CBP[23] result in the loss of this inhibition. Similarly, stimulation of CBP HAT activity was seen with mutants of pCIP that interact with the C-terminus of CBP via the NCBD, but not by the TF CREB that binds to the N-terminal KIX domain of CBP[23] which overlaps CBP_{IDR3} and CBP_{IDR4}. Although more studies using full-length CBP are required to test this hypothesis, the ability to influence CBP HAT activity by altering the balance of condensation mediated by CBP_{IDR6} and CBP_{IDR7} is potentially important for generating specific activity profiles, especially given the prevalence of IDR-IDR interactions and condensate formation among diverse TFs [27,56–58].

By focusing on CBP_{wt} behaviour in the nucleus, we revealed that the AL, which is sufficient to drive condensate formation of $p300_{HAT}$ *in vitro*[22], is not the primary driver of full-length CBP_{wt} condensates. Instead, the AL positively contributes to CBP_{wt} condensates within an overall hierarchy (Figure 6B), but also provides a platform for negative regulation via lysine acetylation (Figure 4B-E). Notably, p300-dependent lysine acetylation can reduce chromatin association mediated by C-terminal TF interacting domains in p300 [11], consistent with negative regulation of condensate size and intensity by lysine acetylation (Figure 4D) [22] and the ability of CBP condensates to change patterns of chromatin localisation (Figure 5F). The lysine-rich AL is also an RNA-binding region, facilitating RNA-dependent regulation of HAT activity[19]. As CBP_{AL} can modulate the properties of condensates, this raises the possibility

that binding of RNA to the AL could also regulate CBP condensate properties, analogous to the effect of RNA on condensates formed by NONO[59].

As the sequence determinants of condensate formation become better understood, more studies are revealing how IDRs can work together to modulate condensate behaviour. Here we highlight how cooperation between CBP_{IDR6} and CBP_{IDR7} facilitates variability in the number and behaviour of CBP condensates (Figure 6B). The inhibitory effect of CBP_{IDR6} on condensates driven by CBP_{IDR7} (Figure 6C) is similar to regulatory interactions in condensates formed by Ataxin-2, where condensate forming IDRs - with prion-like properties similar to CBP_{IDR7} - could be quenched by neighbouring IDRs[60]. Our results highlight the importance of fine-grained dissection of IDRs for uncovering the complexity inherent in the condensate behaviour of large regulatory proteins: although studies of p300 had previously implicated the disordered C-terminus as a driver of condensates[9], the larger-scale deletions did not uncover the opposing contributions of individual IDRs on condensate properties.

Stabilisation of transient secondary structures from conformational ensembles is a functional mechanism of IDRs[26]. CBP_{IDR6} and CBP_{IDR7} bridge the NCBD of CBP (Figure 4I & S4G), a TF binding domain that exists in a molten globule state and which undergoes ‘synergistic folding’ to form helical bundles with specific topologies when bound to different TFs[18,35,36,61,62]. Both CBP_{IDR6} and CBP_{IDR7} overlap the NCBD, raising the possibility that disrupted NCBD folding and subsequent changes to TF binding could underlie changes in condensate behaviour we observed here. However, we note that deletion of a region equivalent to the NCBD in p300 had only mild effects on p300 chromatin association[11]. The molten globule conformation adopted by the NCBD and the high propensity of regions in both CBP_{IDR6} and CBP_{IDR7} to form helical structures (Figure S4E-F) is a reminder that while the entire C-terminal region of CBP is largely disordered, these regions are likely sampling a number of transient secondary structure conformations. While we did not observe disruption of CBP condensates upon deletion of helices that are predicted to flank the NCBD (Figure S4H-J, localised helical structures in disordered regions can tune the properties of condensates in other systems[63], and there remains the possibility that TF binding to the NCBD could alter the balance of interactions between CBP_{IDR6} and CBP_{IDR7} and thus change the properties of CBP condensates.

CBP and its paralog p300 are often thought of interchangeably, due to the high sequence identity of the catalytic domain. However, CBP and p300 each perform distinct functions in cells and *in vitro* [64–66], although the molecular basis of these distinct functions is poorly understood. Notably, the sequence identity between CBP and p300 is considerably lower in their C-terminal IDRs important for regulating the behaviour of CBP condensates, than in

the well conserved HAT domain (Figure S4D). Our data highlights how subtle differences in the sequence properties and behaviour of CBP_{IDR6} and CBP_{IDR7} can cause pronounced differences in the overall properties of CBP condensates. We note that the sequence properties of these regions are different again in p300 (Figure4 J-O), raising the intriguing possibility that p300 condensates may behave differently to CBP condensates, with corresponding implications for their biological function. Interestingly, $p300_{wt}$ was earlier found to have a lower ability to form condensates than individual large p300-IDRs, suggesting that uncharacterised inhibitory IDR's may be present in p300[9]. While these differences between CBP and p300 are intriguing, extensive further experiments (e.g domain swapping) are needed to fully elucidate the relationship between these IDRs and their effect on p300 condensates.

CBP is recruited to chromatin via interactions with TFs at almost all regulatory elements across the genome[1]. While formation of liquid-like condensates at regulatory elements has emerged as an important mechanism for controlling gene expression, correct regulation of these condensate properties is crucial. The ability of condensates to regulate transcription requires balance: too much condensation downregulates transcription, suggesting that condensates must exist within a window of optimal properties to guarantee accurate transcriptional output[9,44,46,67]. Here, we highlight how different IDRs within CBP can cooperate to help maintain this balance: CBP-IDRs contribute hierarchically to condensates; they are regulated by lysine acetylation and form interaction sites for different TFs and RNAs. The juxtaposition of forces that promote and inhibit condensate formation in adjacent CBP-IDRs appears to form a regulatory module that allows CBP to respond to its nuclear environment by forming condensates with different properties. This potentially facilitates the integration of different external stimuli by CBP, and the fine tuning of its transcription regulatory activities to contribute to enhancer-specific activity profiles. The interactions uncovered provide important new insights into how distinct IDRs contribute to the complex regulation of CBP activity at regulatory elements, and will underpin future studies to further map the biological function of CBP-IDRs.

Acknowledgements

D.B. and AT are Sir Henry Dale Fellows, funded by the Wellcome Trust and the Royal Society (grant numbers 213501/Z/18/Z and 220192/Z/20/Z respectively); K.G., D.B., A.T. and T.D.C. are supported by a BBSRC Pioneer Award (grant number: BB/Y513453/1); K.G. was supported by a studentship from the White Rose BBSRC Doctoral Training Partnership

(grant number: BB/M011151/1); T.D.C was funded by (BB/T008032/1); D.B also received funding from The Royal Society (grant number: RSG\R1\180410). M.D. is part of the Sheffield Bioinformatics Core team. At the University of Sheffield, we thank Susan Clark for support at the Flow Cytometry Core Facility; IT services for access to High Performance Computing and the Wolfson Light Microscopy Facility. The Nikon W1 spinning disk confocal was funded by a BBSRC ALERT2021 award (grant number: BB/V019368/1).

Disclosure and competing interests statement

The authors declare no competing interests.

Data availability

The ChIP-seq data from this publication have been deposited to the GEO database (https://www.ncbi.nlm.nih.gov/geo/) and assigned the identifier GSE268266. All DNA constructs used are available from Addgene.

Methods

Cell culture

Culture and maintenance

Human Embryonic Kidney cells stably expressing the SV40 large T antigen (HEK293T, ATCC) and GP2-293 (Clontech) cells were cultured in complete DMEM (10% FBS, 1% Penicillin-Streptomycin (Gibco)) and grown in 5% CO_2 at 37°C. HEK293T cells containing stable doxycycline inducible CBP-GFP, were cultured in complete DMEM (10% Tetracycline free FBS, 1% Penicillin-Streptomycin) including 2 mg/ml G418 (Generon) and 50ug/ml hygromycin (ThermoFisher) . All cell lines were screened every three months for mycoplasma infection.

Transient transfections

Transient transfections were done using either Lipofectamine 3000 (ThermoFisher Scientific) or LipoD293 (TEBU-Bio). Manufacturers recommended protocol was followed for 6-well dishes with the following details: 1.5 µg of DNA was transfected per well and 7.5 µl of Lipofectamine 3000 reagent was used per well. For LipoD293 (TEBU-Bio), the manufacturers protocol for 6-well 35-mm dishes was followed using: 1 µg of DNA and 3 µl of LipoD293 per well.

Retroviral transduction

Lentiviral transduction of CBP-GFP was done using the pRetroX^{TM} Tet-On^{R} 3G inducible expression system (Clontech). CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP and $CBP_{\Delta CFID}$-GFP were cloned into the pRetroX-pTRE3G vector for lentiviral transduction. To generate lentiviral particles, 2ug of CBP-pRetroX-pTRE3G constructs or pRetroX-Tet3G were co-transfected with 2ug of the envelope pVSV-G plasmid into GP2-293 cells using lipofectamine 3000 (ThermoFisher). After 48 hours, with 1 media change after 24 hours, the viral supernatant was collected and concentrated using Lenti-X concentrator (Takara) following manufacturers instructions. To generate polyclonal pRetroX-Tet3G HEK293T cells, 8 ug/ml polybrene (Merck Millipore) was added the concentrated virus and used to infect HEK293T cells. Infected cells were selected using G418 (Generon) at a concentration of 8 mg/ml for 1 week. To generate CBP-pRetroX-pTRE3G cells, polyclonal pRetroX-Tet3G HEK293T cells were transduced with viral particles and 8 ug/ml polybrene, then selected using 8 mg/ml of G418

and 125ug/ml hygromycin (ThermoFisher) for 1 week. Clonal cell lines were generated by inducing expression of CBP-GFP with 2000 ng/ml of doxycycline for 48 hours before fluorescence-activated cell sorting (FACS) based on the GFP signal, using a FACSMelody Cell Sorter (BD Biosciences) at the University of Sheffield Flow Cytometry Core Facility. Clonal lines were cultured in the background of 2 mg/ml G418 and 50ug/ml hygromycin (ThermoFisher).

Doxycycline induction of CBP expression:

For doxycycline validation experiments, HEK293T cells were plated in 10 cm dishes (Sarstedt) for western blot or 35-mm glass bottom dishes (Ibidi) for imaging. 24 hours post-plating the media was changed to media containing the predetermined concentration of doxycycline calculated to ensure induction of CBP-GFP but not increase overall CBP expression levels (Figure S5A-B)), 2 mg/ml G418 and 50ug/ml hygromycin.

Molecular Biology and cloning

All PCR reactions were completed using a 50 ul reaction containing: 0.5 ul PFuUltra II fusion HS DNA Polymerase (Agilent), 1 % PFu Ultra reaction buffer, 5 % dimethyl sulfoxide (DMSO, Sigma), 200mM dNTP mix, 0.2 uM forward and reverse primer, and template DNA. DNA was initially denatured at 95°C with an extension temperature of 68°C for DNA less than 10 Kb, and 72°C for DNA greater than 10 Kb, with 35 extension cycles unless otherwise stated. To confirm correct assembly and presence of mutations, all plasmids were checked by Sanger sequencing and most were checked by Whole plasmid sequencing (Eurofins). All Constructs have been made available through Addgene (https://www.addgene.org/).

Full length human CBP (CBP_{wt})

To generate human CBP_{wt} (transcript variant 1, NCBI Reference: NM_004380.3) tagged with a C-terminal GFP for overexpression experiments (CBP_{wt}-GFP), *Homo sapiens* CBP transcript variant 2 (NCBI Reference NM_001079846.1, Sino Biological HG17295-UT) was cloned in-frame into pAcGFP-N1 (Clontech) using NEBuilder HiFi DNA Assembly reaction protocol (NEB). CBP was amplified using primers: *Fwd: Human CBP_fwd, Rev: Human CBP_rev* and pAcGFP-N1 was amplified using primers: *Fwd: pAcGFP-N1_fwd, Rev: pAcGFP-N1_rev*; (For primer sequences see Supplementary table 1). Assembled constructs were checked by Sanger sequencing (Eurofins). To convert transcript variant 2 to transcript variant 1, the missing sequence comprising exon 5 was ordered as a custom-synthesised

DNA from IDT and inserted into CBP-pAcGFP-N1 using the NEB HiFi assembly. CBP exon 5 was amplified using primers: *CBP missing Exon_fwd, CBP missing Exon_rev*. CBP-pAcGFP-N1 was amplified using primers: *Fwd: GFP-CBP_fwd, Rev: GFP-CBP_rev*. Fully assembled CBP_{wt}-GFP was sequenced by Sanger sequencing using sequential primers at 900bp intervals and confirmed using Whole Plasmid Sequencing (Eurofins) to check for correct assembly.

CBP-IDRs for optoDroplet

The optoDroplet construct pHR-mCh-Cry2WT was a gift from Clifford Brangwynne (Addgene plasmid #101221)[38]. To target the optoDroplet construct to the nucleus we used custom synthesised DNA (gBlock, IDT) encoding the SV40 nuclear localisation signal (NLS) to generate pHR-mCh-Cry2WT-NLS[34]. The SV40 NLS was amplified by PCR using primers: *Fwd: SV40_NLS_NotIRD_fwd, Rev: SV40_NLS_SbfIRD_rev*, to introduce restriction sites for NotI and SbfI. Following restriction digestion with NotI and SbfI (NEB), pHR-mCh-Cry2WT and the synthesised NLS were ligated using T4 DNA ligase (NEB) and the construct was checked by Sanger sequencing. CBP-IDRs were cloned into pHR-mCh-Cry2WT-NLS construct using the NEB HiFi DNA assembly protocol. pHR-mCh-Cry2WT-NLS was amplified using primers: *Fwd: pHR-mCh-Cry2WT-NLS_fwd, Rev: pHR-mCh-Cry2WT-NLS_rev*. CBP-IDRs were amplified from CBP_{wt}-GFP using IDR-specific primers (Supplementary table 1). All constructs were checked using Sanger sequencing.

$CBP_{\Delta IDR}$-GFP deletion mutants

To generate $CBP_{\Delta IDR}$ mutants, primers were designed to amplify CBP_{wt}-GFP in reverse directions to delete residues comprising the identified CBP-IDRs (Supplementary table 1). Amplified PCR products were digested using DpnI (NEB), the ends of the PCR fragments were then phosphorylated using T4 polynucleotide kinase (PNK, NEB) before undergoing ligation using T4 DNA ligase (NEB). Constructs were sequenced by Sanger sequencing to check the IDRs had been correctly removed from the construct.

Autoregulatory Loop (AL) mutants

Two mutants of the Autoinhibitory Loop (AL, $CBP_{1558-1607}$) were designed. Firstly, CBP_{KTG} contained the AL with all lysines mutated to glycine (K1564G, K1565G, K1583G, K1586G, K1587G, K1588G, K1591G, K1592G, K1595G, K1597G, K1605G, K1606G, and K1607G). Secondly, in $CBP_{\Delta AL}$ the AL was replaced with a glycine/serine linker (sequence

GSAGSAAGSGQF) to maintain correct folding[45]. $CBP_{4322-5818}$ - which are flanked by unique restriction sites for NruI and PmlI - were synthesised as gBlocks (IDT) containing the mutant AL sequences (Supplementary table 1). The synthesised gBlocks and CBP-pAcGFP-N1 constructs (CBP-GFP for single mutants; $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP, and $CBP_{\Delta CFID}$-GFP for double mutants) were digested using NruI and PmlI (NEB). The digested constructs were dephosphorylated using Quick Calf Intestinal Phosphatase (CIP, NEB), before the inserts and vectors were assembled using T4 DNA ligase (NEB). The assembled constructs were checked by Sanger sequencing.

Doxycycline inducible CBP-GFP

CBP-GFP, $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP, and $CBP_{\Delta CFID}$-GFP were sub-cloned into a modified pRetroX-pTRE3G plasmid (Clontech) containing a hygromycin resistance cassette using BamHI-HF and NotI-HF (NEB) restriction enzymes. The fragments were annealed using T4 DNA ligase (NEB) before sequencing to check for correct insertion and subsequently by whole plasmid sequencing.

$CBP_{\Delta IDR6H2}$-GFP and Cloning $CBP_{\Delta IDR7H1}$-GFP:

To delete predicted a-helices in CBP_{IDR6} and CBP_{IDR7}, DNA sequences were synthesised containing CBP 5332 - 7674 which resides within the N-terminal GFP tag, and spans two unique endogenous restriction sites for MluI-HF and BssHII (GeneArt, ThermoFisher). Synthesized sequences contained deletions of $CBP_{1951-1978}$ ($CBP_{\Delta IDR6H2}$-GFP) and $CBP_{2187-2216}$ ($CBP_{\Delta IDR7H1}$-GFP). CBP_{wt}-GFP and synthesised sequences were digested with MluI-HF and BssHII (NEB) and purified by agarose gel extraction. The digested CBP_{wt}-GFP vector was dephosphorylated using Quick Calf Intestinal Phosphatase (CIP, NEB) before assembly using T4 DNA ligase (NEB). Correct insertion was checked by Sanger sequencing.

Cloning of gRNA plasmids

Designed gRNAs (Supplementary table 1) were cloned into pSpCas9(BB)-2A-GFP (PX458). PX458 was a gift from Feng Zhang (Addgene plasmid # 48138) [68]. 100 µM forward and reverse oligo pairs were combined with T4 polynucleotide kinase (PNK) (NEB) and 10X T4 ligation buffer (NEB) containing essential ATP. For annealing, reactions were heated to 37°C for 30 minutes, followed by 95°C for a further 5 minutes and then cooled to 25°C at a ramp rate of -5°C per minute. Annealed gRNAs were then diluted 1:100 in sterile H_2O in preparation for the ligation reaction.

pX458 plasmid was digested with BbsI-HF (NEB) DNA and annealed gRNA oligos were ligated using T4 DNA ligase (NEB). Constructs were validated using Sanger sequencing.

Construction of stably integrated CBP-Halotag cell lines.

gRNA design

Single guide RNAs (sgRNAs) targeting the C-terminus of CBP were designed using the UCSC CRISPR targets tool (https://genome.ucsc.edu/). After selection in UCSC, crRNAs were further evaluated using the Evaluation tool on the E-CRISP site (http://www.e-crisp.org/E-CRISP/reannotate_crispr. html). Plasmids encoding CBP-targeting sgRNAs were transfected into HEK293T cells using the Neon electroporation system and 100 µl kit (ThermoFisher), at 1100V, 20 ms pulse width and 2 pulses. After 48 hours, genomic DNA was extracted using QuickExtract solution (Lucigen) according to the manufacturer's instructions and PCR of CBP C-terminal region was carried out using primers *CBP_C-T_fwd1/2, CBP_C-T_rvs* (Supplementary Table 1).

gRNA validation

PCR products amplifying the C-terminal region of CBP in both edited and unedited cells were cleaned up using the Monarch PCR cleanup kit (NEB). DNA concentrations were measured using the Qubit dsDNA BR kit (ThermoFisher) according to the manufacturer's instructions. 10X NEB buffer 2 (NEB) was added to 200 ng DNA and PCR products were denatured to single strands by heating at 95°C for 5 minutes. DNA strands were then hybridised by decreasing temperature from 95°C to 85°C at a ramp rate of -2°C/second, then from 85°C to 25°C at a ramp rate of -0.1^{o}C per second. T7 Endonuclease 1 (T7E1) enzyme (NEB) was added and incubated at 37°C for 15 minutes, followed by quenching with 0.25 M EDTA. The total reaction volume was run on a 1% agarose gel. To analyse Sanger sequencing data of CRISPR edited DNA sequences, .ab1 files of CBP C-terminal PCR products were uploaded to the Inference of CRISPR Edits (ICE, Synthego) tool (https://ice.synthego.com/#/, Conant et al. 2022).

Generation of clonal CBP-Halo HEK293T cell lines:

Validated plasmids encoding CBP gRNA and custom synthesised CBP HDR template plasmid (Supplementary Table 1, ThermoFisher) were co-transfected into HEK293T cells using the Neon electroporation system and 100 µl kit (ThermoFisher) at 1100V, 20 ms pulse

width and 2 pulses. After 48 hours, single cell sorting was performed using a FACSMelody cell sorter (BD Biosciences) into 96 well plates. Clones were allowed to grow until confluent and then screened PCR genotyping for successful integration of the Halotag using primers *CBP_C-T_fwd1/2, CBP_C-T_rvs* (Supplementary Table 1) .

qPCR

Total RNA was extracted from WT and CBP-HaloTag HEK293T cells, using TRIzol (ThermoFisher) and treated with Turbo DNase (ThermoFisher) to remove genomic DNA, then purified using phenol/chloroform and ethanol precipitation. Reverse Transcription was performed using the High Capacity cDNA kit (ThermoFisher) and qPCR was carried out using Power SYBR reaction mastermix (ThermoFisher) and a Quantstudio 12K Flex (ThermoFisher). A mean Ct was calculated using each triplicate reaction and normalised to 18S rRNA expression.

Imaging

Live cell imaging of CBP-HaloTag

HEK293T cells containing endogenous CBP-HaloTag were seeded at 400,000 cells/well on 35 mm glass bottom dishes that have been PLL coated (as previously described). After 48 hours, cells were labelled with either HaloTag ligand TMR (Promega) at a concentration of 50 nM for 15 minutes at 37ºC in the dark. Cells were left for a further 2 hours before staining with Hoechst 33342 (Fisher) in DPBS as described in a previous section, before washing in DPBS and storing in imaging media (FluoroBright DMEM (Fisher) supplemented with a final concentration of 20 nM Gibco HEPES (ThermoFisher Scientific)) immediately prior to imaging. Cells were imaged using the Nikon W1 spinning disc confocal microscope equipped with a temperature stage set to 37ºC. Images were taken using the 100X oil immersion lens (NA 1.45), using the 405 nm laser to image Hoechst using 45 % laser power with a 100 ms exposure, TMR was visualised using the 514 nm laser using 60 % laser power with an exposure of 300 ms. Images were taken every 500 ms in the absence of binning. Captured images and videos were processed ImageJ/Fiji [69].

FRAP of CBP-HaloTag

CBP-HaloTag cells were plated on PLL coated 35-mm dishes, and labelled with Janelia Fluor® 549 (JF549, Promega) at a concentration of 50 nM for 15 minutes. Cells were

subsequently washed twice in DPBS and placed in imaging media (as above) for a 15 Yeminute recovery. Cells were imaged using a custom built scanning TIRF/FRAP imaging system (Cairn) with Nikon Ti2-E inverted stand, Cairn multiline laserbank, iLas2 unit and Photometrics Prime 95B sCMOS camera. Imaging was performed at 37ºC in a humified enclosure (OKO Lab). Puncta were imaged in near-TIRF using 561 nm laser with 200 ms exposure at 10% laser power, using a Nikon CFI Apochromat 100X oil NA 1.49 objective. ROIs were drawn with a diameter of 1.1 µm (10 pixels), with multiple puncta per field of view. FRAP settings are as follows: 10 repetitions using the 561 nm laser line of weight 2 on 3 % laser power. A time course was taken with the following pattern: 15 pre-bleach frames at 2 fps; FRAP (performed as described above); 60 recovery frames at 2 fps; 30 frame at 0.5 fps. Images were quantified and processed in ImageJ/Fiji [69].

Imaging of transfected CBP-HaloTag with CBP-GFP, $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP

Endogenous CBP-HaloTag cells were seeded at 250,000 cells/well on PLL coated glass bottom 35-mm dishes and transfected after 24 hours with CBP-GFP, $CBP_{\Delta IDR6}$-GFP, $CBP_{\Delta IDR7}$-GFP using LipoD293, as has been previously described. 24 hours after transfection cells were labelled with Janelia Fluor® 646 (JF646, Promega) at a final concentration of 2 nM for 15 minutes before washing in DPBS and imaging in imaging media. 35 mm dishes were labelled sequentially, immediately before imaging. 4-colour control images were also taken where the HaloTag had not been labelled, to check for bleed-through of overexpressed constructs. In this case, the media was changed to imaging media and imaged using laser lines 405 nm, 488 nm, 532 nm and 637 nm, with an exposure time of 200 ms using 10 % laser power. A 4-colour control was also taken with CBP-HaloTag labelled with JF646 and no overexpression, for this all laser lines used the same conditions as before except the 637 nm laser had a power of 80 % to visualise the endogenous protein. For colocalization experiments, 2 colour multidimensional images were captured using the 488 nm laser with 10 % laser power and a 200 ms exposure to visualise GFP tagged constructs, and 637 nm laser with 80 % laser power and 200 ms exposure to view endogenously tagged HaloTag. Images were processed in ImageJ/Fiji [69].

Imaging of doxycycline inducible clones:

For imaging cells were seeded at a concentration of 250,00 cells/well on PLL (poly-L-lysine) coated 35-mm dishes and doxycycline induced for 48-hours. Imaging was performed in imaging media using the Nikon W1 spinning disc confocal microscope. Laser lines 405 nm and 488 nm were used to image the Hoechst 33342 (Fisher) and therefore the nuclei, and

GFP expression respectively. The 405 nm laser was set with a 100 ms exposure with 60 % laser power, and the 488 nm laser was set with 300 ms exposure time and 60 % laser power. For quantification Z-stacks were taken to capture the whole volume of the nuclei in 0.5 µm increments, these were further processed in ImageJ/Fiji [69]. For the plus doxycycline conditions, imaging media was supplemented with the appropriate concentration of doxycycline.

optoDroplet

35-mm glass bottomed dishes (Ibidi) were coated in PLL for 1 hour at 37°C before being washed twice with dH_2O and seeded with HEK293T cells at a concentration of 200,000 cells/well. 24 hours post-plating the cells were transfected using LipoD293™(SignaGen Laboratories), following manufacturer's guidelines for 35 mm dishes. After 24 hours the media was changed into complete DMEM (10% FBS, 1% Pen/Strep). After a further 24 hours, prior to imaging the media was changed into FluoroBright DMEM (Fisher) with a final concentration of 20 mM HEPES (ThermoFisher). Live cell imaging was performed on Nikon A1 confocal using the 60X oil immersion objective (NA 1.4), equipped with a humidified temperature stage set to 37°C. Two laser wavelengths were used to induce global activation of the optoDroplet system; 488 nm to induce the dimerisation of the Cry2 and 560 nm to image mCherry. An image was captured every 8 seconds, with subsequent activation of the mCherry for the same time. Time courses were taken for each construct, for a period of up to 3 minutes. 3 time courses were taken per construct per biological replicate, with a total of 3 biological replicates being taken.

Fluorescence recovery after photobleaching (FRAP) in transfected cells.

For FRAP experiments, 250,000 cells/well of HEK293T cells were plated in 35 mm dishes and transfected using LipoD293 as was described for the optoDroplet live cell imaging. 24-hours post transfection the media was removed and the cells were washed in DPBS before leaving the cells in imaging media. Imaging was performed using the custom built scanning TIRF/FRAP imaging system (Cairn) described above using a Nikon CFI Apochromat 100X oil NA 1.49 objective. Puncta were visualised in near-TIRF using the 488 nm laser set to 5 % laser power with a 100 ms exposure. A circular ROI with dimensions of 4.4 x 4.4 µm was drawn over the puncta to be photobleached; there may be multiple regions within each field of view. Photobleaching was performed using the 488 nm laser, with 10 % laser power, using 5 repetitions with a thickness of 5. 10 images were taken prior to FRAP, followed by 700 frames being collected after, with a frame rate of ~2.6 fps.

1,6 - Hexanediol treatment:

Cells were plated and transfected as previously described for FRAP in transfected cells. Images were captured using the Nikon W1 spinning disc confocal microscope using the 100X oil immersion objective. Images were taken using the 488 nm laser with a 200 ms exposure time on 20 % laser power. A 10% solution of 1,6 - Hexanediol (Merck) was diluted from a 50% stock in warm imaging media. Time courses were taken of the constructs over a 3 minute period, imaging every 500 ms. Without pausing imaging, after 30 seconds 1 ml of 10% 1,6 - Hexanediol was added dropwise to the 35 mm dish that contained 1 ml of imaging media for a final concentration of 5% 1,6 - Hexanediol. Imaging continued until the end of the time course. Image quantification and processing was performed in ImageJ/Fiji [69].

Immunofluorescence and fixed cell imaging

Immunofluorescence and fixed cell imaging

Circular coverslips (Ibidi) were acid washed in 0.25% acetic acid before washing and storing in 100% ethanol. Coverslips were coated in poly-L lysine (PLL, Merck) for 1 hour at 37°C before washing twice in sterile dH_2O. HEK293T cells were grown in 6-well plates at a concentration of 200,000 cells/well containing the coverslips for 48 hours before fixing. If the cells were transfected, transfection was performed 24 hours after plating using Lipofectamine 3000 transfection reagent (ThermoFisher) or LipoD293 (TEBU-Bio) following manufacturers instructions. Cells were washed in DPBS twice before fixing in 2% paraformaldehyde (PFA) for 30 minutes at room temperature in the dark before being washed 3 times in DPBS. For immunofluorescence, coverslips were blocked using a blocking buffer (5% Bovine serum albumin (BSA) with 0.2% Triton in PBS) for 1 hour at 37°C. Primary antibody (for full list of antibodies see Supplementary Table 2) was prepared in a blocking buffer and incubated at 37°C for 1 hour. After washing 3 times in PBS, secondary antibodies were diluted in the blocking buffer and incubated for 37°C for 1 hour. After fixing, coverslips were mounted onto slides (VWR, 631-1554) using VECTORshield that contained 4,6-diamidino-2-phenylindole (DAPI) (Vector laboratories, H-2000-10). Fixed slides were stored at 4°C. Imaging work was performed at the Wolfson Light Microscopy Facility using the Nikon A1 confocal microscope. All fixed cell imaging was performed using the 60X oil immersion objective (NA 1.4). The microscope has 4 laser wavelengths, typically only 3 of the 4 were used (these include, 405 nm for DAPI, 488 nm for GFP, 562 nm for mCherry and 642 nm for Alexa Fluor 647). Z-stacks were taken every 0.5 µm to cover the

depth of the nuclei; a minimum of 3 fields of view were imaged per construct over 2 coverslips, per biological replicate.

Image analysis

Integrated Intensity Density

Before calculating the Integrated density, individual nuclei were isolated and using a blinded approach, the number of nuclei that display either puncta or a diffuse phenotype was determined for each of the constructs. From this a percentage of nuclei which show either puncta or diffuse was calculated.

Integrated density, the intensity multiplied by the area, for each condensate produced by a defined construct was calculated in ImageJ/Fiji [69] using a custom macro. A maximum projection was generated from each Z-stack, and the condensates were identified by applying a manual intensity threshold to each image to generate a mask. The intensity threshold was internally consistent within each experimental repeat, where the threshold was chosen to best isolate condensates across the range of different phenotypes observed within the constructs tested. Within the mask, each independent ROI was labelled, and the area, Integrated Density, and circularity was measured for each ROI.

Once these images had been analysed, the percentage of nuclei which contained a ROI which has an intensity higher than that of the set intensity threshold value was further calculated. This value was described as the transfection efficiency, and was generated to identify which nuclei had passed the threshold value. Filters were then applied sequentially to the data: 1) To exclude overexpressed nuclei, a maximum grey value of 10 was set; 2) To remove regions of high intensity that were too small to represent puncta, an area threshold (<3 pixels/0.36um) was applied; 3) To remove regions of overlapping puncta, a circularity threshold (<0.25) was set. All downstream analysis was done using R in Rstudio; significance p values for comparison of CBP-GFP behaviours (number of condensate per nuclei and integrated density) were calculated using a Kruskal-Wallis test. Graphs and statistical testing was performed using R in RStudio [70].

optoDroplet

A custom ImageJ macro was written to identify puncta formed in the mCherry channel. Images were run through an ImageJ plugin called Stack Registration (StackReg, [71]), and a threshold was applied using RenyiEntropy to generate a binary mask that underwent binary

processing of closing followed by opening. The parameters were then set in analyse particles, applying a size range of 50 - infinity, to generate ROIs which equated to the outline of the signal region, which was the nuclei. These ROIs were then applied to the original, raw image that underwent a multi-measure to calculate the standard deviation of the pixels within the nuclei over the time course. Regions found in the first frame were used to monitor the change in the standard deviation over the time course. Average Fold Change was then calculated, as well as rate of change over the first 32 seconds of the time course. The statistics for the rate of change were performed in GraphPad prism, where an unpaired t-test was used to determine statistical significance (Prism [72]). Graphs for visualisation were made using R in RStudio [70].

FRAP in CBP-HaloTag

Images for each FRAP time course were loaded into ImageJ/Fiji [69] and processed using StackReg ([71]) to remove drift in the sample. Over each time course, the intensity was calculated for a region of interest around the bleached area, before being corrected for photobleaching over the time course using a reference region within the same nuclei. Relative intensity was calculated by subtracting the intensity at T = 0, where, this was defined as the first frame after photobleaching, and then normalised to the pre-bleach intensity for each construct. Rate constants and T = ½ were calculated in GraphPad Prism 9 [72], using a non-linear regression with settings one phase association with a variance weighting of $1/Y^2$.

FRAP in transfected cells

Time courses from FRAP of overexpressed protein were quantified using ImageJ/Fiji [69].The intensity was calculated over the time course, using a ROI of 0.22 µm, which is smaller than the condensate to only monitor the intensity within the condensate. The condensate was manually tracked throughout the time course, moving the ROI when the condensate moved. The relative fluorescence recovery was calculated by subtracting the intensity at T = 0, the first frame post bleaching and then normalised to the pre-bleach intensity for each construct. Rate constants and T = ½ were calculated as described for FRAP of CBP-HaloTag.

Western blotting

HEK293T cells were harvested and after washing twice in PBS the pellet was lysed in RIPA buffer (50 mM Tris-HCl pH 8 stored at 4°C, 100 mM NaCl, 2 mM $MgCl_2$, 1 % Triton X-100, 0.1 % Sodium deoxycholate, 0.1 % Sodium Dodecyl Sulfate (SDS)) supplemented with 1

mM Dithiothreitol (DTT), 1X Halt protease inhibitor cocktail, 10 mM sodium butyrate and 500 U/µl benzonase. 25 - 50 ug of overall protein containing sample buffer and reducing buffer was loaded onto a 3 - 8% tris-acetate gel (Thermo Fisher Scientific) using tris acetate running buffer (Life Technologies). Gel was transferred onto a membrane using trans-blot turbo mini 0.2 µm nitrocellulose transfer packs (Biorad) before blocking and probing for CBP, GFP and GAPDH (for full list of antibodies see Supplementary Table 2). Western blots were imaged using G:box (Syngene) gel imager. Images were processed using ImageJ/Fiji [69] and assembled using Adobe Illustrator.

Chromatin Immunoprecipitation (ChIP):

For ChIP experiments, Dox inducible cells were plated in 10 cm dishes, 2 plates per construct. Cells were induced with concentrations of Dox that were identified using western blot as inducing CBP-GFP without altering overall CBP expression levels (Figure S5A-B). . After 48h, cells were crosslinked at room temperature for 10 minutes using formaldehyde (1.1 % final). Crosslinking was quenched using 2.5 M Glycine (Melford, 125 mM final) incubating for 5 minutes at room temperature. Cells were harvested by cell scraping in PBS, and pelleted by centrifugation at 200 rcf for 5 minutes. Cell lysis was performed in ChIP lysis buffer 1 (50 mM HEPES-KOH (Sigma), pH 7.5 140 mM NaCl, 1 mM EDTA, 10 % Glycerol, 0.5 % NP40, 0.25 % Triton X - 100, Complete protease inhibitor cocktail (Roche)) by rotating at 4°C for 5 minutes. Nuclei were pelleted by centrifugation at 1500 rcf for 5 minutes, re-suspended in ChIP Buffer 2 (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA (Sigma), Complete protease inhibitor cocktail) and incubated at room temperature on a rotator for 10 minutes. Nuclei were pelleted by centrifugation at 1500 rcf for 5 minutes and re-suspended in ChIP lysis buffer 3 (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1 % Na-deoxycholate, 0.5 % N-lauroylsarcosine (Sigma), Complete protease inhibitor cocktail) for shearing. Chromatin was sheared to 150 - 300 bp fragments, using the Bioruptor Pico (Diagenode), samples were sonicated for 12 cycles of 30 seconds on, 30 seconds off pulse sonication. Lysates were cleared by centrifugation at 20,000xg for 20 min. Concentrations of lysates was determined by BCA assay (ThermoFisher), IP's were done with 900 µg of sheared protein in a total of 300 µl with ChIP IP buffer (10 mM Tris-HCl pH 8.0, 600 mM NaCl, 1 mM EDTA, 3 % Triton X-100, Complete protease inhibitor cocktail), input reactions were made for each sample with 90 µg of protein in 50 µl of ChIP IP buffer. Immunoprecipitations were performed by incubating with 4 µg of antibody overnight with rotation at 4°C. Immunocomplexes were recovered by adding blocked Protein G Dynabeads and incubated for 90 minutes at 4°C with rotation. Beads were washed 5 times in ChIP wash

buffer (50 mM HEPES-KOH pH 7.5, 500 mM LiCl, 1 mM EDTA, 1 % NP40, 0.7 % Na-deoxycholate, 0.1 % N-lauroylsarcosine) and 1x in ChIP final wash buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 50 mM NaCl). Sample was eluted by incubation at 65°C for 30 min in ChIP elution buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 10 mM EDTA, 1 % SDS) with rotation in a thermomixer at 900 rpm. Cross-linking was reversed by incubation at 65ºC overnight. Immunoprecipitated DNA was treated with RNase A (0.2 mg/ml final), supplemented with 4 mM $CaCl_2$ for 2 hours at 37ºC, then Proteinase K (0.2 mg/ml final) for 2 hours at 55ºC. DNA was then purified by phenol:chloroform extraction and ethanol precipitation and resuspended in 1x TE buffer.

ChIP-seq

For ChIP-seq, samples were prepared using NEBNext Ultra™II DNA Library Prep Kit for Illumina (NEB, E7645) following manufacturers instructions. Briefly, fragmented DNA underwent 5' phosphorylation and dA-Tailing to prepare the fragment ends for adapter ligation. Sample concentration was checked using Qubit dsDNA High Sensitivity (HS) kit (Thermo Fisher), to determine adapter dilutions. For input samples the adapters were not diluted, however a 1:25 dilution of adapters in 10 mM Tris-HCl pH 7.5, was performed for immunoprecipitation samples. A cleanup of the adapter ligation was performed using AMPure XP beads (Beckman Coulter) for size selection following the 200 bp protocol. PCR amplification was performed with 3 cycles of amplification for input samples, and 13 cycles of amplification for immunoprecipitation samples. A cleanup of the PCR reaction was performed using AMPure XP beads (Beckman Coulter), following protocol guidelines. Fragment sizes were checked before sequencing using the TapeStation (Agilent), using High Sensitivity DNA ScreenTape (Agilent).

ChIP-seq Data analysis:

ChIP-seq samples were sequenced using paired-end 150 sequencing on an Illumina NovaSeq 6000 platform (Novogene). Raw data quality was assessed using FastQC v0.11.9[73] and adapters were trimmed using Cutadapt v3.4 [74] implemented in Trim Galore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Trimmed reads were aligned to human genome assembly GRCh38 using Burrows-Wheeler Alignment (BWA) v0.7.17-r1188[75]. For downstream analysis, replicates were merged and duplicate reads removed using Picard v2.27.4 (https://broadinstitute.github.io/picard/), before read quality filtering using SAMTools v1.15.1[76]. BigWig files were generated by normalizing per million reads mapped (RPM) and converting to bedgraph format with Bedtools v2.30.0[77], then to BigWig format using UCSC utilities (bedGraphToBigWig v377)[78]. ChIP-seq peaks were

called on filtered data using MACS2 v2.2.7.1[79] using the BAMPE option and a broad peak cutoff of 0.001. Called peaks were assigned to annotated genome regions using Homer software (http://homer.ucsd.edu/homer/ngs/annotation.html). All initial processing, alignment and peak calling was completed as part of the Nextflow ChIP-seq pipeline v2.00[80] on the University of Sheffield High Performance Computing (HPC) cluster. Intersection of ChIP-seq peaks, calculation of relative distance plots[54] and Jaccard statistics[54] were done using Bedtools v2.31.0. Control peaks corresponding to each sample were calculated using bedtools shuffle, over a window restricted to 40kb upstream of transcription start sites on the same chromosome. Subsequent analysis was done using a local installation of DeepTools v3.5.5 [81]. Read density coverage was calculated using multiBigwigSummary in 1kb bins over the entire genome; Correlation coefficients between samples were calculated using plotCorrelation; Heatmaps and metaplots displaying signals over a 5kb window aligned to peak centres were generated using computeMatrix and plotHeatmap; Read density enrichment over defined peak regions were calculated using plotEnrichment. For enrichment of H3K27ac over CBP peaks, we used publicly available data: H3K27ac ChIP-seq in wt HEK293T cells (GEO: GSM1249889)(Liu et al. 2013). All downstream analysis was done using R in Rstudio; significance p values for tag enrichment were calculated using a two-sided Mann-Whitney U-test in R[70]. The ChIP-seq data from this publication have been deposited to the GEO database (https://www.ncbi.nlm.nih.gov/geo/) and assigned the identifier GSE268266.

Computational analysis

Identification of CBP-IDRs

To predict disordered regions in CBP, we used DISOPRED3 [82] and PONDR[33] disorder prediction software. Tested CBP-IDRs were determined using PONDR VSL2 predictor [33], where stretches of amino acids that have a prediction of disorder above 50% were classed as intrinsically disordered regions [34]. A mutant was also generated spanning the region where CBP binds to FUS [37] we termed the CBP-FUS Interaction Domain (CFID).

Analysis of CBP sequence properties

Analysis of sequence properties was carried out using a local installation of localCIDER[43]. Linear composition plots for full length CBP_{wt} and CBP_{CFID} were calculated using a blob

length window of 200bp and 50bp respectively. PLAAC analysis was carried out using the web-based PLAAC server[51].

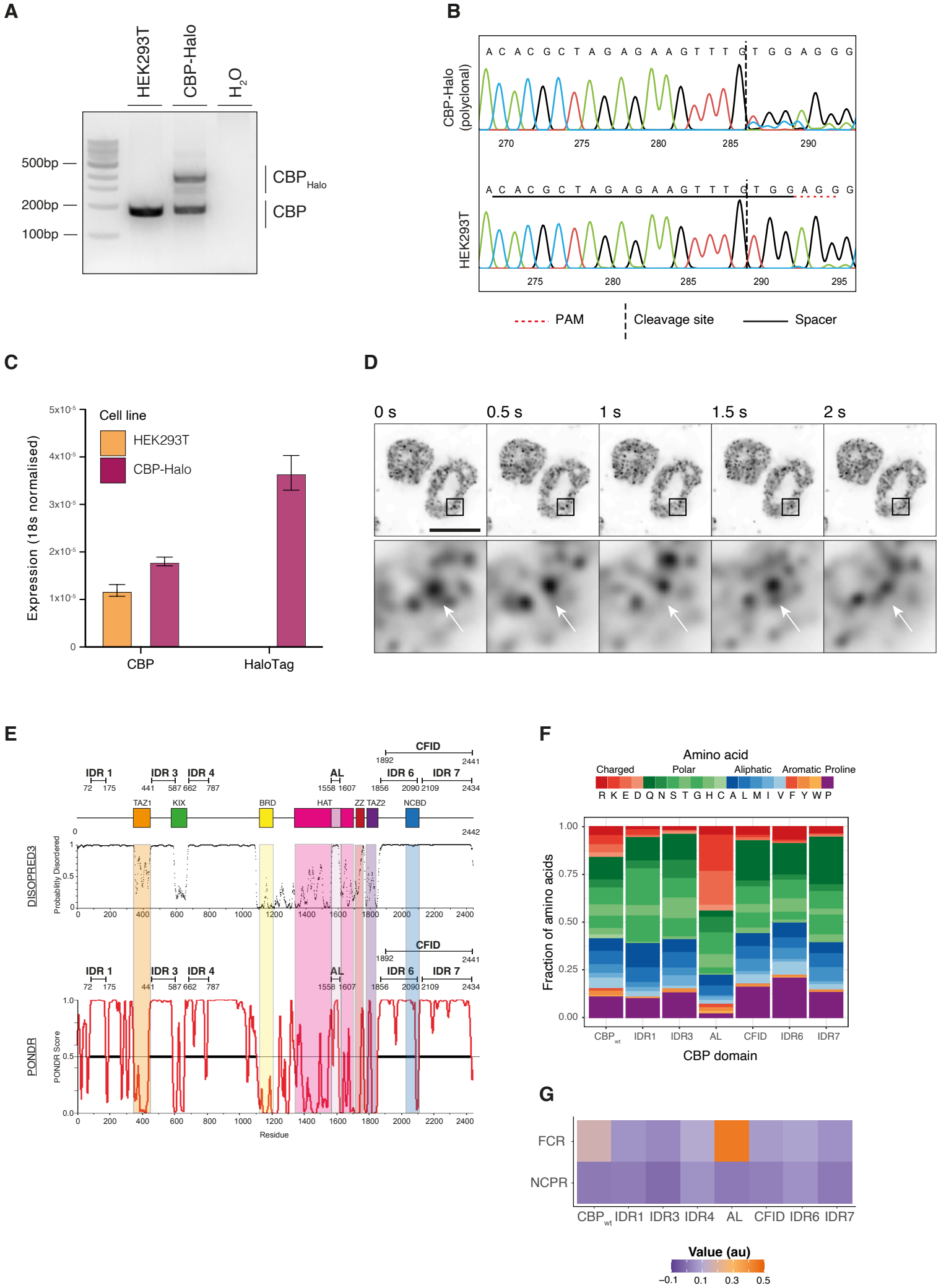
AlphaFold structure prediction (version 2.3.2)

All structure predictions were performed using AlphaFold version 2.3.2[32] implemented through a Colab notebook. We used the monomer model with relaxation and specified five recycles. AlphaFold produces a per-residue confidence metric called predicted local distance difference test (pLDDT) on a scale from 0 to 100. pLDDT estimates how well the prediction would agree with an experimental structure; a pLDDT >70 is considered to correspond to a generally correct backbone prediction[83]. To assess the confidence of predictions carried out here, we provided a duplicate view of our model coloured by pLDDT score (Figure S4F). All visualisation was done using UCSF ChimeraX[84].

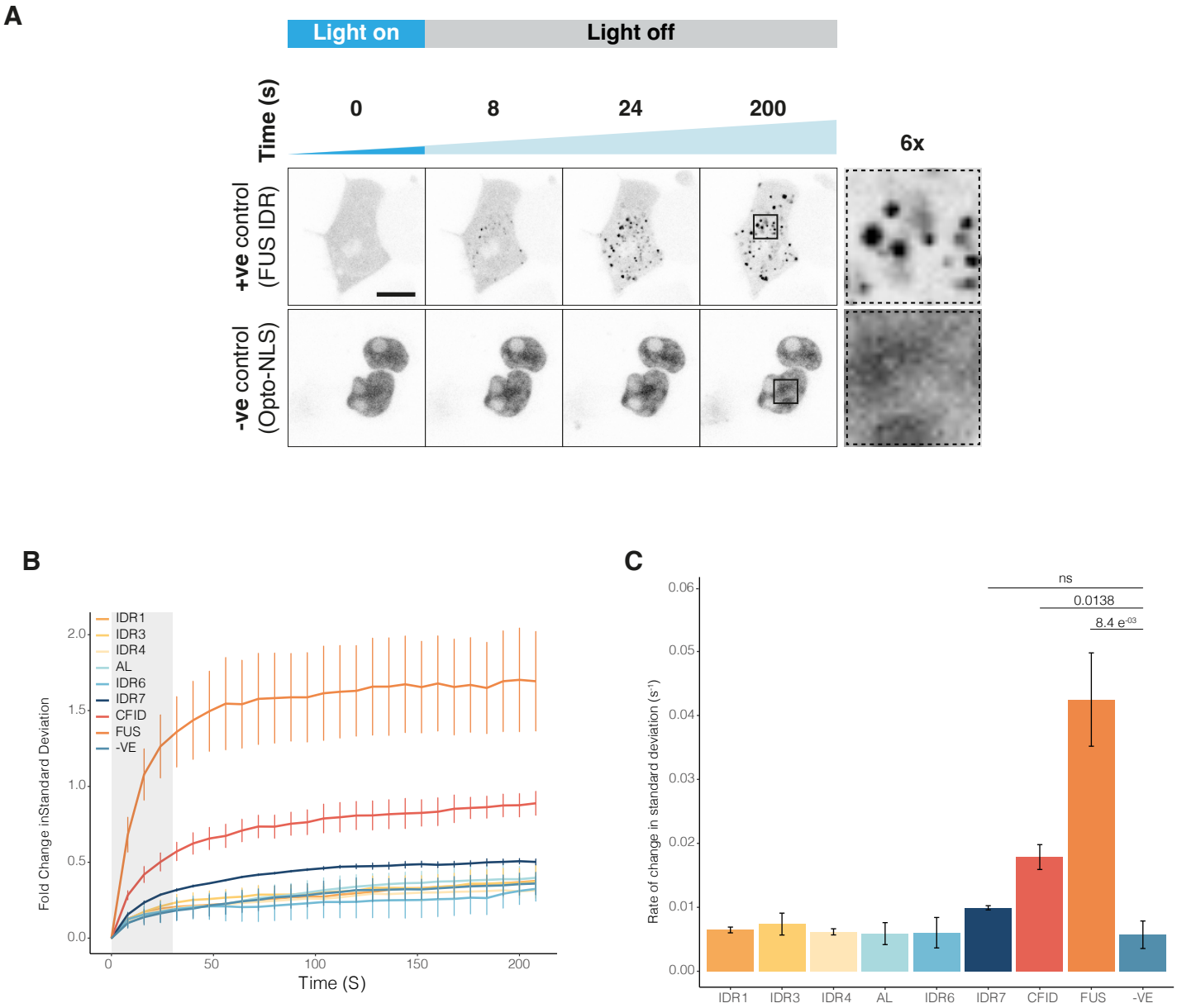
Data availability

The ChIP-seq data from this publication have been deposited to the GEO database (https://www.ncbi.nlm.nih.gov/geo/) and assigned the identifier GSE268266. All DNA constructs used are available from Addgene.

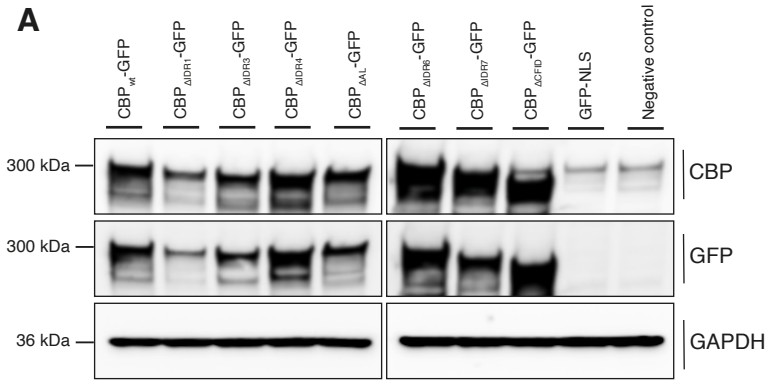
Supplementary Figure 1

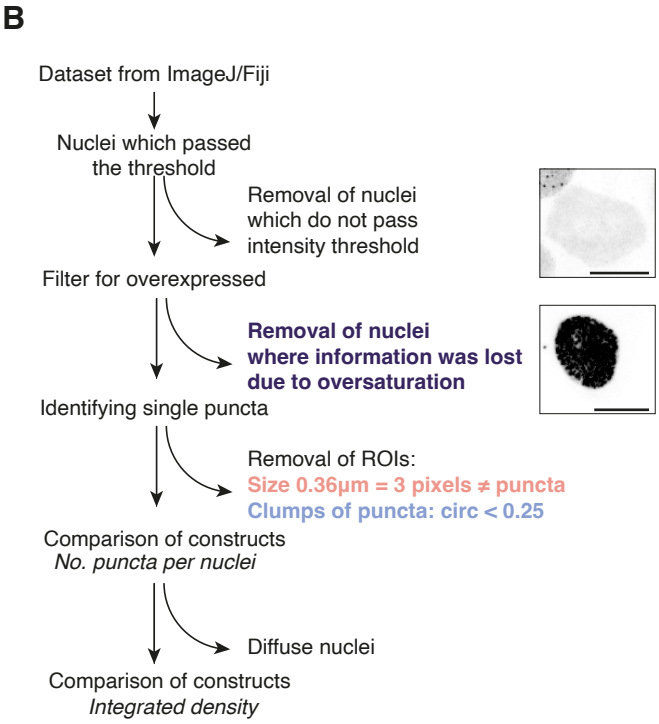


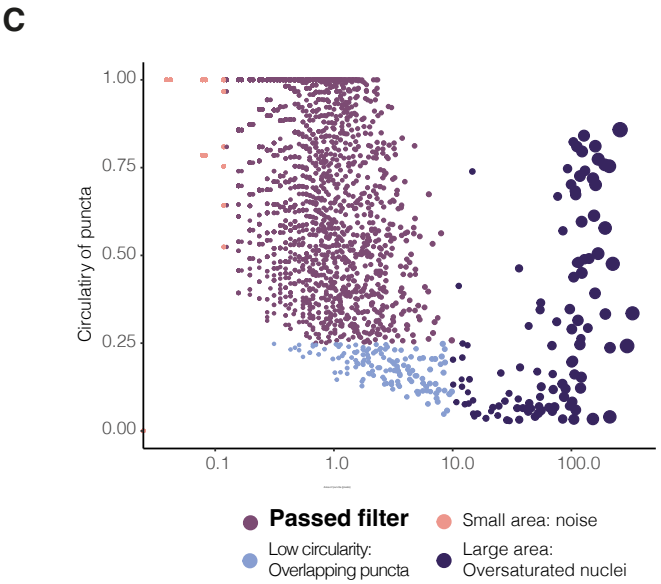
Supplementary Figure 2

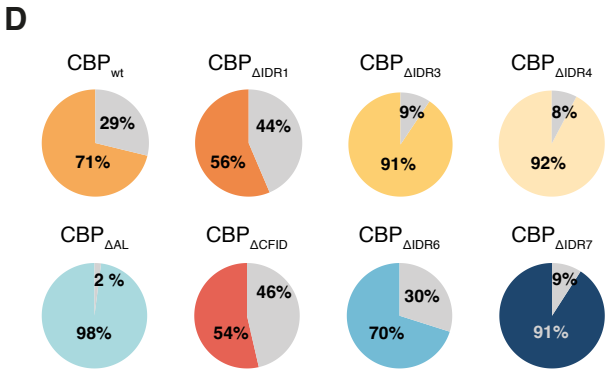


Supplementary Figure 3

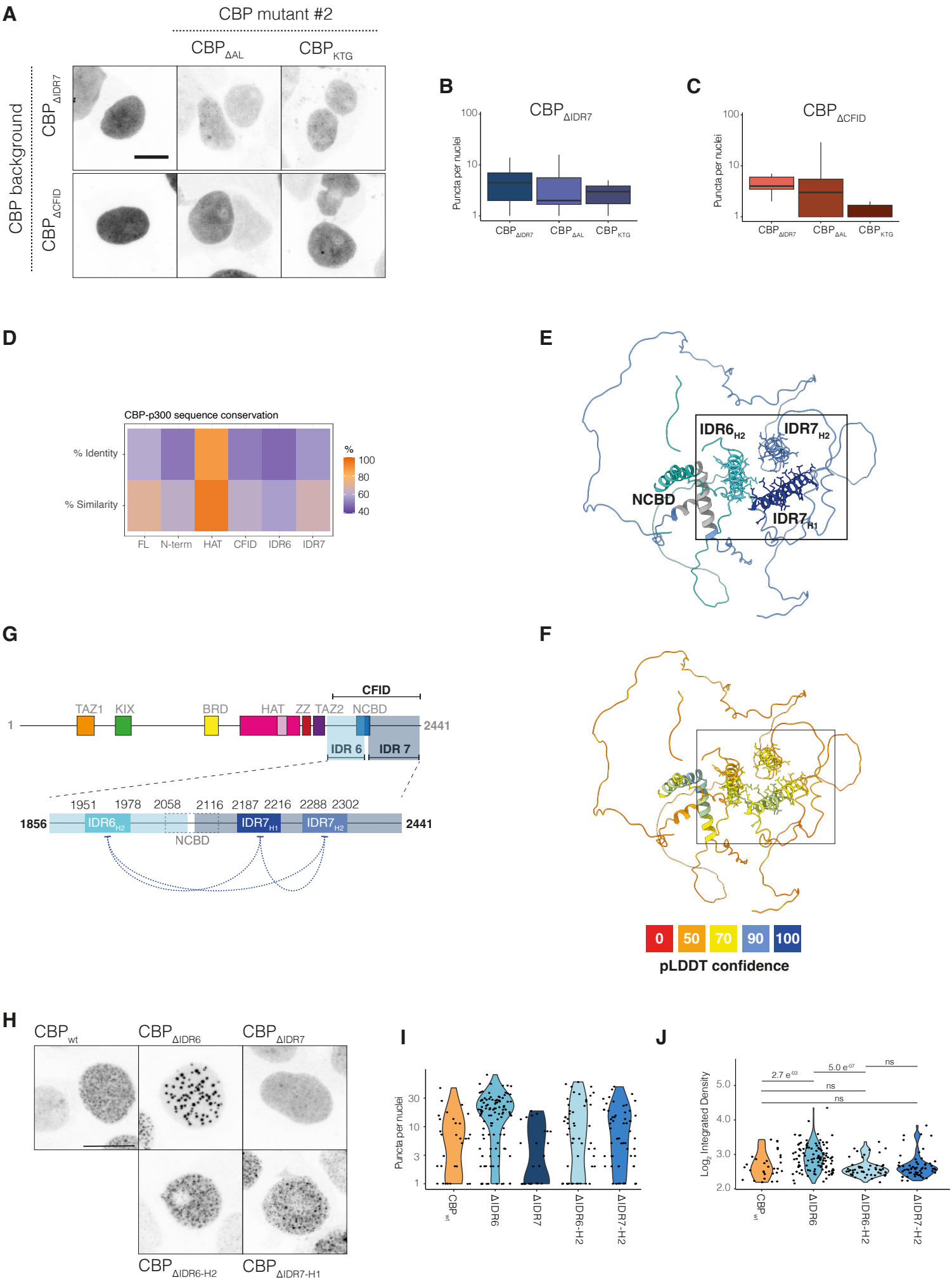




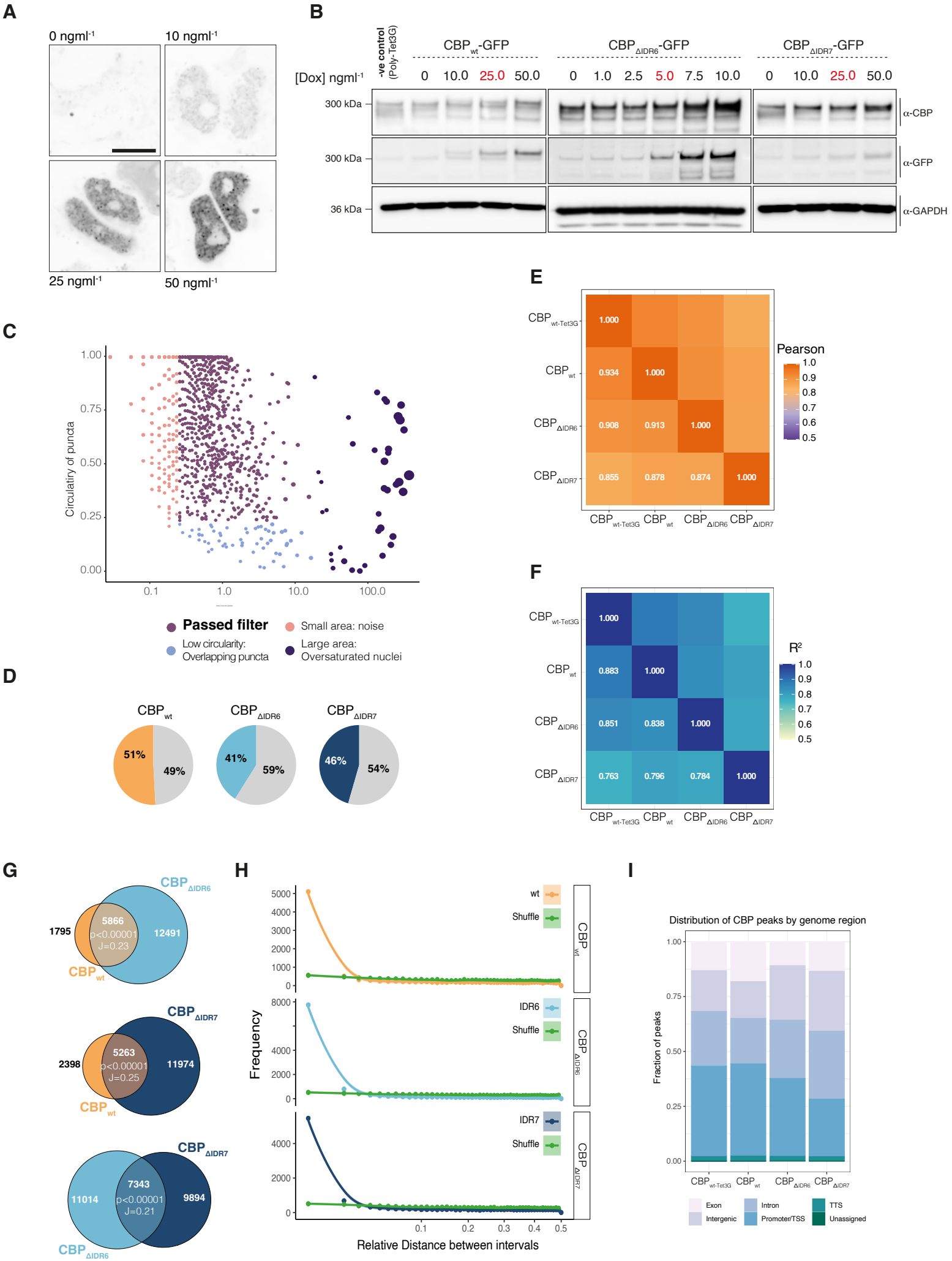




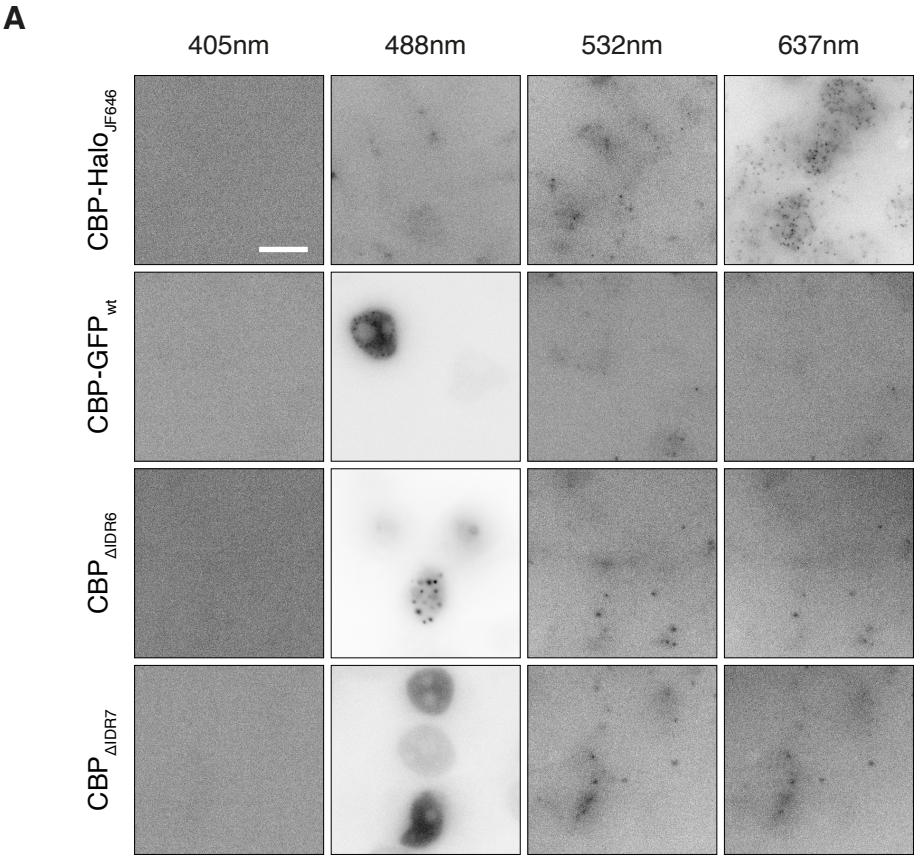
Supplementary Figure 4



Supplementary Figure 5



Supplementary Figure 6



Supplementary Figure Legends

Figure S1: Introduction of C-terminal HaloTag to endogenous CBP.

A-B) Analysis of PCR products from edited locus: (A) PCR using primers spanning the 3' end of CBP_{wt}; (B) Inference of CRISPR Edits (ICE)[86] analysis of Sanger sequencing data from polyclonal CBP-Halo HEK293T (Top) and HEK293T control cells (bottom).

C) RT-qPCR using primers for CBP or HaloTag in control HEK293T cells (orange) and CBP-Halo HEK293T cells (magenta). Bars represent mean ± s.e.m, n=3.

D) Still images from Supplementary Movie 1 showing the dynamic behaviour of CBP-Halo condensates - Alternative representative condensates from (Figure 1D). Boxes highlight regions enlarged below; arrows highlight representative dynamic condensates; Scale bars: 10 µm.

E) Domain map and predicted intrinsic disorder of full length CBP_{wt} using: DISOPRED3[82], Top; PONDR[33], bottom. Positions of CBP-IDRs (CBP_{IDR1-7}, CBP_{AL} and CBP_{CFID}) with their domain boundaries predicted using PONDR, and structured CBP domains are highlighted.

F-G) Sequence properties for CBP_{wt} and CBP-IDRs calculated using localCIDER[43]. E) Amino acid composition of CBP_{wt} and tested CBP-IDRs; F) Fraction of Charged Residues (FCR, top) and Net Charge per Residue (NCPR, bottom).

Figure S2 CBP IDRs have distinct properties and behaviours

A) Confocal microscopy images showing behaviour of positive control (FUS-IDR) and negative control (mCherry-Cry2-NLS, Opto-NLS) in the OptoDroplet assay. Cells were imaged for 200s following exposure to blue light. Scale bars: 10 µm.

B) Enlarged view of 200s time point from Figure 2C highlighting clear puncta for CBP_{IDR7} and CBP_{CFID}. Scale bars: 10 µm.

C) Fold change in the standard deviation of the grey values for defined individual nuclei following blue light stimulation. Diffuse signal should have a low standard deviation, while strong changes in intensity due to condensate formation should have a high standard deviation. Grey shading highlights the 30 second window used to calculate the initial rate of change. Data same as in Figure 2F, but also includes positive control (FUS-IDR). Bars represent mean ± s.e.m; significant p values are reported (ns, not significant). n=3.

D) Initial rate of change over the first 30 seconds following blue light exposure. Data same as in Figure 2F, but also includes positive control (FUS-IDR). Bars represent mean ± s.e.m; significant p values are reported (ns, not significant), n=3.

Figure S3: CBP-IDRs affect the behaviour of full length CBP_{wt}

A) Western blot for CBP and GFP shows relative expression levels of transiently transfected CBP-GFP constructs.

B) Filtering strategy used for selecting transfected nuclei for analysis. Transfected nuclei were selected based on an intensity threshold [69]. GFP puncta were then filtered by: i) Area - to remove overexpression artefacts (e.g oversaturated fluorescent nuclei); ii) noise less than 3 pixels in diameter; iii) Circularity - to isolate single from overlapping puncta. Images show examples of nuclei removed by the filter, scale bars: 10 µm.

C) Summary of filtering results. Highlighted populations are: Nuclei passing filter (Purple); Large area to remove oversaturated cells/nuclei with no clear puncta (dark blue); low circularity (<0.25) to remove overlapping puncta (light blue); Puncta < 3 pixels to remove speckle noise (salmon).

D) Percentage of transfected nuclei for each condition remaining following application of the filter.

Figure S4: Contribution of CBP_{AL} to CBP condensates and sequence properties of C-terminal regions.

A) Fixed cell confocal microscopy of CBP_{AL} double mutant constructs. Scale bars: 10 µm.

B-C) Mean number of condensates per nuclei formed by $CBP_{\Delta AL}$-GFP and CBP_{KTG}-GFP compared to background mutant. B) $CBP_{\Delta IDR7}$-GFP background; C) $CBP_{\Delta CFID}$-GFP

D-E) AlphaFold v2.3.2 [32] prediction of the isolated CBP_{CFID}. D) The positions the predicted interacting α-helices CBP_{IDR6H2} (cyan), CBP_{IDR7H1} (dark blue), CBP_{IDR7H2} (light blue), and the NCBD are highlighted; E) AlphaFold v2.3.2[32] prediction of CBP_{CFID} highlighting pLDDT confidence scores, same view as in D).

F) Domain structure of the CBP_{CFID} highlighting the positions of CBP_{IDR6} (cyan) and CBP_{IDR7} (light blue). The NCBD and predicted helices are highlighted. Dashed lines show the interactions between α-helices predicted by AlphaFold v2.3.2 [32].

G) Fixed cell confocal microscopy comparing the behaviour of CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP with $CBP_{\Delta IDR6H2}$-GFP and $CBP_{\Delta IDR7H1}$-GFP. Scale bars: 10 µm.

H) Amino acid sequence conservation between CBP and p300 from local pairwise sequence alignment using EMBOSS Matcher[85]. N-term, regions N-terminal to the HAT domain residues CBP 1-1195; p300 1-1159. (Full alignments are available in Supplementary material).

I-J) Comparison of condensates formed by $CBP_{\Delta IDR6H2}$-GFP and $CBP_{\Delta IDR7H1}$-GFP with CBP_{wt}-GFP, $CBP_{\Delta IDR6}$-GFP and $CBP_{\Delta IDR7}$-GFP. I) Number of puncta per nuclei; J) Integrated intensity density. p values were calculated using a Kruskal-Wallis test.

Figure S5: Chromatin localisation of inducible CBP-GFP.

A-B) Inducible CBP-GFP expression. CBP_{wt}-GFP HEK293T cells were treated with 0-50ng/ml Doxycycline for 48h. A) Confocal microscopy images. Scale bars: X µm; B) Western blot for CBP and GFP. The concentration of Dox that induced CBP-GFP expression while maintaining consistent total selected for induction is highlighted (red).

C-D) Summary of filtering results. C) Highlighted populations are: Nuclei passing filter (Purple); Large area to remove oversaturated cells/nuclei with no clear puncta (dark blue); low circularity (<0.25) to remove overlapping puncta (light blue); Puncta < 3 pixels to remove speckle noise (salmon); D) Percentage of nuclei which express CBP-GFP or CBP-GFP mutant constructs following the application of the filter.

E-F) Correlation between ChIPseq datasets. D) Pearson's correlation coefficient; E) R^2.

G) Venn diagrams showing overlap between GFP Peaks in CBP-GFP cell lines; CBP_{wt}-GFP (orange), $CBP_{\Delta IDR6}$-GFP (light blue) and CBP_{IDR7}-GFP (dark blue). p values from permutation test with random regions restricted to TSS -40 kb; Jaccard statistic (J) measures overlap between two sets of intervals; perfect correlation = 1.

H) Relative distance plots for CBP-GFP ChIPseq peaks (GFP ChIP). When intervals are closer than expected by chance, distributions shift towards lower relative distance values. Shuffled control regions (green) were restricted to TSS -40 kb.

I) Distribution of CBP localisation by genome region.

Figure S6: $CBP_{\Delta IDR}$ mutants affect endogenous CBP condensates

A) Multiple channel control for live cell imaging. CBP-Halo was labelled with Janelia Fluor® 646 Halo ligand (CBP-$Halo_{JF646}$).

References

1. Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A., Boyer, L. A., Young, R. A. & Jaenisch, R. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 21931–21936 (2010).
2. Tie, F., Banerjee, R., Stratton, C. A., Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M. O., Scacheri, P. C. & Harte, P. J. CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. *Development* **136,** 3131–3141 (2009).
3. Tropberger, P., Pott, S., Keller, C., Kamieniarz-Gdula, K., Caron, M., Richter, F., Li, G., Mittler, G., Liu, E. T., Bühler, M., Margueron, R. & Schneider, R. Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer. *Cell* **152,** 859–872 (2013).
4. Pradeepa, M. M., Grimes, G. R., Kumar, Y., Olley, G., Taylor, G. C. A., Schneider, R. & Bickmore, W. A. Histone H3 globular domain acetylation identifies a new class of enhancers. *Nat. Genet.* **48,** 681–686 (2016).
5. Jin, Q., Yu, L.-R., Wang, L., Zhang, Z., Kasper, L. H., Lee, J.-E., Wang, C., Brindle, P. K., Dent, S. Y. R. & Ge, K. Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO J.* **30,** 249–262 (2011).
6. Weinert, B. T., Narita, T., Satpathy, S., Srinivasan, B., Hansen, B. K., Schölz, C.,

Hamilton, W. B., Zucconi, B. E., Wang, W. W., Liu, W. R., Brickman, J. M., Kesicki, E. A., Lai, A., Bromberg, K. D., Cole, P. A. & Choudhary, C. Time-resolved analysis reveals rapid dynamics and broad scope of the CBP/p300 acetylome. *Cell* **174,** 231–244.e12 (2018).

7. Stasevich, T. J., Hayashi-Takanaka, Y., Sato, Y., Maehara, K., Ohkawa, Y., Sakata-Sogawa, K., Tokunaga, M., Nagase, T., Nozaki, N., McNally, J. G. & Kimura, H. Regulation of RNA polymerase II activation by histone acetylation in single living cells. *Nature* (2014). doi:10.1038/nature13714
8. Gomez-Lamarca, M. J., Falo-Sanjuan, J., Stojnic, R., Abdul Rehman, S., Muresan, L., Jones, M. L., Pillidge, Z., Cerda-Moya, G., Yuan, Z., Baloul, S., Valenti, P., Bystricky, K., Payre, F., O'Holleran, K., Kovall, R. & Bray, S. J. Activation of the notch signaling pathway in vivo elicits changes in CSL nuclear dynamics. *Dev. Cell* **44,** 611–623.e7 (2018).
9. Ma, L., Gao, Z., Wu, J., Zhong, B., Xie, Y., Huang, W. & Lin, Y. Co-condensation between transcription factor and coactivator p300 modulates transcriptional bursting kinetics. *Mol. Cell* **81,** 1682–1697.e7 (2021).
10. Narita, T., Ito, S., Higashijima, Y., Chu, W. K., Neumann, K., Walter, J., Satpathy, S., Liebner, T., Hamilton, W. B., Maskey, E., Prus, G., Shibata, M., Iesmantavicius, V., Brickman, J. M., Anastassiadis, K., Koseki, H. & Choudhary, C. Enhancers are activated by p300/CBP activity-dependent PIC assembly, RNAPII recruitment, and pause release. *Mol. Cell* (2021). doi:10.1016/j.molcel.2021.03.008
11. Ferrie, J. J., Karr, J. P., Graham, T. G. W., Dailey, G. M., Zhang, G., Tjian, R. & Darzacq, X. p300 is an obligate integrator of combinatorial transcription factor inputs. *Mol. Cell* **84,** 234–243.e4 (2024).
12. Taylor, G., Eskeland, R., Hekimoglu-Balkan, B., Pradeepa, M. & Bickmore, W. A. H4K16 acetylation marks active genes and enhancers of embryonic stem cells, but does not alter chromatin compaction. *Genome Res.* (2013). at <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=23990607&re

tmode=ref&cmd=prlinks>

13. Kasper, L. H., Qu, C., Obenauer, J. C., McGoldrick, D. J. & Brindle, P. K. Genome-wide and single-cell analyses reveal a context dependent relationship between CBP recruitment and gene expression. *Nucleic Acids Res.* **42,** 11363–11382 (2014).
14. Wang, Z., Chivu, A. G., Choate, L. A., Rice, E. J., Miller, D. C., Chu, T., Chou, S.-P., Kingsley, N. B., Petersen, J. L., Finno, C. J., Bellone, R. R., Antczak, D. F., Lis, J. T. & Danko, C. G. Prediction of histone post-translational modification patterns based on nascent transcription data. *Nat. Genet.* **54,** 295–305 (2022).
15. Hunt, G., Boija, A. & Mannervik, M. p300/CBP sustains Polycomb silencing by non-enzymatic functions. *Mol. Cell* **82,** 3580–3597.e9 (2022).
16. Boija, A., Mahat, D. B., Zare, A., Holmqvist, P.-H., Philip, P., Meyers, D. J., Cole, P. A., Lis, J. T., Stenberg, P. & Mannervik, M. CBP Regulates Recruitment and Release of Promoter-Proximal RNA Polymerase II. *Mol. Cell* **68,** 491–503.e5 (2017).
17. Bedford, D. C. & Brindle, P. K. Is histone acetylation the most important physiological function for CBP and p300? *Aging* (2012). at <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=22511639&retmode=ref&cmd=prlinks>
18. Dyson, H. J. & Wright, P. E. Role of Intrinsic Protein Disorder in the Function and Interactions of the Transcriptional Coactivators CREB-binding Protein (CBP) and p300. *J. Biol. Chem.* **291,** 6714–6722 (2016).
19. Bose, D. A., Donahue, G., Reinberg, D., Shiekhattar, R., Bonasio, R. & Berger, S. L. RNA Binding to CBP Stimulates Histone Acetylation and Transcription. *Cell* **168,** 135–149.e22 (2017).
20. Carullo, N. V. N. & Day, J. J. Genomic Enhancers in Brain Health and Disease. *Genes* **10,** (2019).
21. Thompson, P. R., Kurooka, H., Nakatani, Y. & Cole, P. A. Transcriptional coactivator protein p300. Kinetic characterization of its histone acetyltransferase activity. *J. Biol. Chem.* **276,** 33721–33729 (2001).

22. Zhang, Y., Brown, K., Yu, Y., Ibrahim, Z., Zandian, M., Xuan, H., Ingersoll, S., Lee, T., Ebmeier, C. C., Liu, J., Panne, D., Shi, X., Ren, X. & Kutateladze, T. G. Nuclear condensates of p300 formed though the structured catalytic core can act as a storage pool of p300 with reduced HAT activity. *Nat. Commun.* **12,** 4618 (2021).
23. Perissi, V., Dasen, J. S., Kurokawa, R., Wang, Z., Korzus, E., Rose, D. W., Glass, C. K. & Rosenfeld, M. G. Factor-specific modulation of CREB-binding protein acetyltransferase activity. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 3652–3657 (1999).
24. Moses, D., Guadalupe, K., Yu, F., Flores, E., Perez, A. R., McAnelly, R., Shamoon, N. M., Kaur, G., Cuevas-Zepeda, E., Merg, A. D., Martin, E. W., Holehouse, A. S. & Sukenik, S. Structural biases in disordered proteins are prevalent in the cell. *Nat. Struct. Mol. Biol.* (2024). doi:10.1038/s41594-023-01148-8
25. Cuevas-Velazquez, C. L., Vellosillo, T., Guadalupe, K., Schmidt, H. B., Yu, F., Moses, D., Brophy, J. A. N., Cosio-Acosta, D., Das, A., Wang, L., Jones, A. M., Covarrubias, A. A., Sukenik, S. & Dinneny, J. R. Intrinsically disordered protein biosensor tracks the physical-chemical effects of osmotic stress on cells. *Nat. Commun.* **12,** 5438 (2021).
26. Holehouse, A. S. & Kragelund, B. B. The molecular basis for cellular function of intrinsically disordered protein regions. *Nat. Rev. Mol. Cell Biol.* 1–25 (2023). doi:10.1038/s41580-023-00673-0
27. Ferrie, J. J., Karr, J. P., Tjian, R. & Darzacq, X. ‘Structure’-function relationships in eukaryotic transcription factors: The role of intrinsically disordered regions in gene regulation. *Mol. Cell* (2022). doi:10.1016/j.molcel.2022.09.021
28. LaMorte, V. J., Dyck, J. A., Ochs, R. L. & Evans, R. M. Localization of nascent RNA and CREB binding protein with the PML-containing nuclear body. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 4991–4996 (1998).
29. Girdwood, D., Bumpass, D., Vaughan, O. A., Thain, A., Anderson, L. A., Snowden, A. W., Garcia-Wilson, E., Perkins, N. D. & Hay, R. T. P300 transcriptional repression is mediated by SUMO modification. *Mol. Cell* **11,** 1043–1054 (2003).
30. Boisvert, F. M., Kruhlak, M. J., Box, A. K., Hendzel, M. J. & Bazett-Jones, D. P. The

transcription coactivator CBP is a dynamic component of the promyelocytic leukemia nuclear body. *J. Cell Biol.* **152,** 1099–1106 (2001).

31. Nucifora, F. C., Sasaki, M., Peters, M. F., Huang, H., Cooper, J. K., Yamada, M., Takahashi, H., Tsuji, S., Troncoso, J., Dawson, V. L., Dawson, T. M. & Ross, C. A. Interference by huntingtin and atrophin-1 with cbp-mediated transcription leading to cellular toxicity. *Science* **291,** 2423–2428 (2001).
32. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P. & Hassabis, D. Highly accurate protein structure prediction with AlphaFold. *Nature* **596,** 583–589 (2021).
33. Xue, B., Dunbrack, R. L., Williams, R. W., Dunker, A. K. & Uversky, V. N. PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochim. Biophys. Acta* **1804,** 996–1010 (2010).
34. Basu, S., Mackowiak, S. D., Niskanen, H., Knezevic, D., Asimi, V., Grosswendt, S., Geertsema, H., Ali, S., Jerković, I., Ewers, H., Mundlos, S., Meissner, A., Ibrahim, D. M. & Hnisz, D. Unblending of Transcriptional Condensates in Human Repeat Expansion Disease. *Cell* (2020). doi:10.1016/j.cell.2020.04.018
35. Demarest, S. J., Martinez-Yamout, M., Chung, J., Chen, H., Xu, W., Dyson, H. J., Evans, R. M. & Wright, P. E. Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415,** 549–553 (2002).
36. Lee, C. W., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Structure of the p53 transactivation domain in complex with the nuclear receptor coactivator binding domain of CREB binding protein. *Biochemistry* **49,** 9964–9971 (2010).
37. Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., Tempst, P., Rosenfeld, M. G., Glass, C. K. & Kurokawa, R. Induced ncRNAs allosterically modify RNA-binding

proteins in cis to inhibit transcription. *Nature* **454,** 126–130 (2008).

38. Shin, Y., Berry, J., Pannucci, N., Haataja, M. P., Toettcher, J. E. & Brangwynne, C. P. Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets. *Cell* **168,** 159–171.e14 (2017).
39. Luo, F., Gui, X., Zhou, H., Gu, J., Li, Y., Liu, X., Zhao, M., Li, D., Li, X. & Liu, C. Atomic structures of FUS LC domain segments reveal bases for reversible amyloid fibril formation. *Nat. Struct. Mol. Biol.* **25,** 341–346 (2018).
40. Patel, A., Lee, H. O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M. Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T. M., Pozniakovski, A., Poser, I., Maghelli, N., Royer, L. A., Weigert, M., Myers, E. W., Grill, S., Drechsel, D., Hyman, A. A. & Alberti, S. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **162,** 1066–1077 (2015).
41. Thompson, P. R., Wang, D., Wang, L., Fulco, M., Pediconi, N., Zhang, D., An, W., Ge, Q., Roeder, R. G., Wong, J., Levrero, M., Sartorelli, V., Cotter, R. J. & Cole, P. A. Regulation of the p300 HAT domain via a novel activation loop. *Nat. Struct. Mol. Biol.* **11,** 308–315 (2004).
42. Park, S., Stanfield, R. L., Martinez-Yamout, M. A., Dyson, H. J., Wilson, I. A. & Wright, P. E. Role of the CBP catalytic core in intramolecular SUMOylation and control of histone H3 acetylation. *Proceedings of the National Academy of Sciences* **114,** E5335–E5342 (2017).
43. Holehouse, A. S., Das, R. K., Ahad, J. N., Richardson, M. O. G. & Pappu, R. V. CIDER: Resources to analyze sequence-ensemble relationships of intrinsically disordered proteins. *Biophys. J.* **112,** 16–21 (2017).
44. Chong, S., Graham, T. G. W., Dugast-Darzacq, C., Dailey, G. M., Darzacq, X. & Tjian, R. Tuning levels of low-complexity domain interactions to modulate endogenous oncogenic transcription. *Mol. Cell* **82,** 2084–2097.e5 (2022).
45. Delvecchio, M., Gaucher, J., Aguilar-Gurrieri, C., Ortega, E. & Panne, D. Structure of the p300 catalytic core and implications for chromatin targeting and HAT regulation.

Nat. Struct. Mol. Biol. (2013). doi:10.1038/nsmb.2642

46. Chong, S., Dugast-Darzacq, C., Liu, Z., Dong, P., Dailey, G. M., Cattoglio, C., Heckert, A., Banala, S., Lavis, L., Darzacq, X. & Tjian, R. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361,** eaar2555 (2018).
47. Ortega, E., Rengachari, S., Ibrahim, Z., Hoghoughi, N., Gaucher, J., Holehouse, A. S., Khochbin, S. & Panne, D. Transcription factor dimerization activates the p300 acetyltransferase. *Nature* **562,** 538–544 (2018).
48. Martin, E. W., Holehouse, A. S., Grace, C. R., Hughes, A., Pappu, R. V. & Mittag, T. Sequence Determinants of the Conformational Properties of an Intrinsically Disordered Protein Prior to and upon Multisite Phosphorylation. *J. Am. Chem. Soc.* **138,** 15323–15335 (2016).
49. Das, R. K. & Pappu, R. V. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 13392–13397 (2013).
50. Alberti, S., Halfmann, R., King, O., Kapila, A. & Lindquist, S. A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. *Cell* **137,** 146–158 (2009).
51. Lancaster, A. K., Nutter-Upham, A., Lindquist, S. & King, O. D. PLAAC: a web and command-line application to identify proteins with prion-like amino acid composition. *Bioinformatics* **30,** 2501–2502 (2014).
52. Lee, C., Kalmar, L., Xue, B., Tompa, P., Daughdrill, G. W., Uversky, V. N. & Han, K.-H. Contribution of proline to the pre-structuring tendency of transient helical secondary structure elements in intrinsically disordered proteins. *Biochim. Biophys. Acta* **1840,** 993–1003 (2014).
53. Piai, A., Calçada, E. O., Tarenzi, T., Grande, A. D., Varadi, M., Tompa, P., Felli, I. C. & Pierattelli, R. Just a Flexible Linker? The Structural and Dynamic Properties of CBP-ID4 Revealed by NMR Spectroscopy. *Biophys. J.* **110,** 372–381 (2016).

54. Favorov, A., Mularoni, L., Cope, L. M., Medvedeva, Y., Mironov, A. A., Makeev, V. J. & Wheelan, S. J. Exploring massive, genome scale datasets with the GenometriCorr package. *PLoS Comput. Biol.* **8,** e1002529 (2012).

55. Liu, W., Ma, Q., Wong, K., Li, W., Ohgi, K., Zhang, J., Aggarwal, A. & Rosenfeld, M. G. Brd4 and JMJD6-associated anti-pause enhancers in regulation of transcriptional pause release. *Cell* **155,** 1581–1595 (2013).

56. Patil, A., Strom, A. R., Paulo, J. A., Collings, C. K., Ruff, K. M., Shinn, M. K., Sankar, A., Cervantes, K. S., Wauer, T., St Laurent, J. D., Xu, G., Becker, L. A., Gygi, S. P., Pappu, R. V., Brangwynne, C. P. & Kadoch, C. A disordered region controls cBAF activity via condensation and partner recruitment. *Cell* **186,** 4936–4955.e26 (2023).

57. Lafontaine, D. L. J., Riback, J. A., Bascetin, R. & Brangwynne, C. P. The nucleolus as a multiphase liquid condensate. *Nat. Rev. Mol. Cell Biol.* **22,** 165–182 (2021).

58. Oksuz, O., Henninger, J. E., Warneford-Thomson, R., Zheng, M. M., Erb, H., Vancura, A., Overholt, K. J., Hawken, S. W., Banani, S. F., Lauman, R., Reich, L. N., Robertson, A. L., Hannett, N. M., Lee, T. I., Zon, L. I., Bonasio, R. & Young, R. A. Transcription factors interact with RNA to regulate genes. *Mol. Cell* **83,** 2449–2463.e13 (2023).

59. Zhang, S., Cooper, J. A. L., Chong, Y. S., Naveed, A., Mayoh, C., Jayatilleke, N., Liu, T., Amos, S., Kobelke, S., Marshall, A. C., Meers, O., Choi, Y. S., Bond, C. S. & Fox, A. H. NONO enhances mRNA processing of super-enhancer-associated GATA2 and HAND2 genes in neuroblastoma. *EMBO Rep.* **24,** (2023).

60. Boeynaems, S., Dorone, Y., Zhuang, Y., Shabardina, V., Huang, G., Marian, A., Kim, G., Sanyal, A., Şen, N.-E., Griffith, D., Docampo, R., Lasker, K., Ruiz-Trillo, I., Auburger, G., Holehouse, A. S., Kabashi, E., Lin, Y. & Gitler, A. D. Poly(A)-binding protein is an ataxin-2 chaperone that regulates biomolecular condensates. *Mol. Cell* (2023). doi:10.1016/j.molcel.2023.05.025

61. Kjaergaard, M., Andersen, L., Nielsen, L. D. & Teilum, K. A folded excited state of ligand-free nuclear coactivator binding domain (NCBD) underlies plasticity in ligand recognition. *Biochemistry* **52,** 1686–1693 (2013).

62. Lin, C. H., Hare, B. J., Wagner, G., Harrison, S. C., Maniatis, T. & Fraenkel, E. A small domain of CBP/p300 binds diverse proteins: solution structure and functional studies. *Mol. Cell* **8,** 581–590 (2001).

63. Conicella, A. E., Dignon, G. L., Zerze, G. H., Schmidt, H. B., D'Ordine, A. M., Kim, Y. C., Rohatgi, R., Ayala, Y. M., Mittal, J. & Fawzi, N. L. TDP-43 α-helical structure tunes liquid–liquid phase separation and function. *Proc. Natl. Acad. Sci. U. S. A.* **117,** 5883–5894 (2020).

64. Martire, S., Nguyen, J., Sundaresan, A. & Banaszynski, L. A. Differential contribution of p300 and CBP to regulatory element acetylation in mESCs. *BMC Mol Cell Biol* **21,** 55 (2020).

65. Henry, R. A., Kuo, Y.-M. & Andrews, A. J. Differences in specificity and selectivity between CBP and p300 acetylation of histone H3 and H3/H4. *Biochemistry* **52,** 5746–5759 (2013).

66. McManus, K. J. & Hendzel, M. J. Quantitative Analysis of CBP- and P300-Induced Histone Acetylations In Vivo Using Native Chromatin. *Molecular and Cellular Biology* **23,** 7611–7627 Preprint at https://doi.org/10.1128/mcb.23.21.7611-7627.2003 (2003)

67. He, J., Huo, X., Pei, G., Jia, Z., Yan, Y., Yu, J., Qu, H., Xie, Y., Yuan, J., Zheng, Y., Hu, Y., Shi, M., You, K., Li, T., Ma, T., Zhang, M. Q., Ding, S., Li, P. & Li, Y. Dual-role transcription factors stabilize intermediate expression levels. *Cell* (2024). doi:10.1016/j.cell.2024.03.023

68. Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A. & Zhang, F. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8,** 2281–2308 (2013).

69. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P. & Cardona, A. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9,** 676–682 (2012).

70. R Core Team. R: A Language and Environment for Statistical Computing. Preprint at https://www.R-project.org/ (2021)

71. Thévenaz, P., Ruttimann, U. E. & Unser, M. A pyramid approach to subpixel registration based on intensity. *IEEE Trans. Image Process.* **7,** 27–41 (1998).

72. San Diego, C. U. *GraphPad Prism 9.0 Software*. (2022). at <www.graphpad.com>

73. Babraham bioinformatics - FastQC A quality control tool for high throughput sequence data. at <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

74. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17,** 10–12 (2011).

75. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25,** 1754–1760 (2009).

76. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. & 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).

77. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26,** 841–842 (2010).

78. Kent, W. J., Zweig, A. S., Barber, G., Hinrichs, A. S. & Karolchik, D. BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* **26,** 2204–2207 (2010).

79. Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W. & Liu, X. S. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9,** R137 (2008).

80. Ewels, P. A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M. U., Di Tommaso, P. & Nahnsen, S. The nf-core framework for community-curated bioinformatics pipelines. *Nat. Biotechnol.* **38,** 276–278 (2020).

81. Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dündar, F. & Manke, T. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44,** W160–5 (2016).

82. Jones, D. T. & Cozzetto, D. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics* **31,** 857–863 (2015).

83. Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., Bridgland, A.,

Cowie, A., Meyer, C., Laydon, A., Velankar, S., Kleywegt, G. J., Bateman, A., Evans, R., Pritzel, A., Figurnov, M., Ronneberger, O., Bates, R., Kohl, S. A. A., Potapenko, A., Ballard, A. J., Romera-Paredes, B., Nikolov, S., Jain, R., Clancy, E., Reiman, D., Petersen, S., Senior, A. W., Kavukcuoglu, K., Birney, E., Kohli, P., Jumper, J. & Hassabis, D. Highly accurate protein structure prediction for the human proteome. *Nature* **596,** 590–596 (2021).

84. Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H. & Ferrin, T. E. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30,** 70–82 (2021).

85. Madeira, F., Pearce, M., Tivey, A. R. N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A. & Lopez, R. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* **50,** W276–W279 (2022).

86. Conant, D., Hsiau, T., Rossi, N., Oki, J., Maures, T., Waite, K., Yang, J., Joshi, S., Kelso, R., Holden, K., Enzmann, B. L. & Stoner, R. Inference of CRISPR edits from Sanger trace data. *CRISPR J.* **5,** 123–130 (2022).